RA
399
U85
1995
v.1

Using Clinical Practice Guidelines To Evaluate Quality of Care

VOLUME 1

ISSUES

Department of Health and Human Services
Donna E. Shalala, Ph.D., *Secretary*

Public Health Service
Philip R. Lee, M.D., *Assistant Secretary for Health*

Agency for Health Care Policy and Research
Clifton R. Gaus, Sc.D., *Administrator*

Office of the Forum for Quality and Effectiveness in Health Care
Douglas Kamerow, M.D., M.P.H., *Director*

NATIONAL INSTITUTES OF HEALTH
NIH LIBRARY
AUG - 2 1995
BLDG 10, 10 CENTER DR.
BETHESDA, MD 20892-1150

Using Clinical Practice Guidelines To Evaluate Quality of Care

Volume 1: Issues

Workgroup Members

Stephen C. Schoenbaum, M.D., M.P.H., Co-chair
David N. Sundwall, M.D., Co-chair
David Bergman, M.D.
June M. Buckle, Sc.D., R.N.
Allan Chernov, M.D.
Janet George, R.N, B.S.N., M.A.
Clark Havighurst, J.D.
M. J. Jurkiewicz, M.D.
John T. Kelly, M.D., Ph.D.
Sandra Metzler, M.B.A., R.R.A.
Christine Miaskowski, Ph.D., R.N., F.A.A.N.
Sam J. W. Romeo, M.D., M.B.A.
Paul M. Schyve, M.D.
Bryan Simmons, M.D.
Patrice Spath, A.R.T., B.A.
Marcia Stevic, Ph.D., R.N.
Constance Winslow, M.D., M.B.A.
Steven Zatz, M.D.

U.S. Department of Health and Human Services
Public Health Service
Agency for Health Care Policy and Research

March 1995

AHCPR Pub. No. 95–0045

PA
397
W85
1995
v.1

Foreword

The legislation that established the Agency for Health Care Policy and Research (AHCPR) in 1989 called for the agency to arrange for the development and periodic review and updating of:

1. Clinically relevant guidelines that may be used by physicians, educators, and health care practitioners to assist in determining how diseases, disorders, and other health conditions can most effectively be prevented, diagnosed, treated, and managed clinically; and

2. Standards of quality, performance measures, and medical review criteria through which health care providers and other appropriate entities may assess or review the provision of health care and assure the quality of such care.

As the process of developing clinical practice guidelines has evolved, the importance of the quality improvement tools that can be derived from guidelines has become clear, as has the need to publish and disseminate a series of reports on using guidelines to improve quality of care. In the applications illustrated in this series of reports, medical review criteria are explicit statements that clinicians, health care organizations, quality and utilization reviewers, payers, and regulators can use to determine how the process of care relates to guidelines. When derived systematically from credible guidelines, review criteria can be used to profile and assess quality of care. Criteria based on comprehensive guidelines can be used to examine how different aspects of the entire care process—diagnosis, treatment, and management—fit together, providing more useful information than isolated rates of incidence, utilization, or occurrence. Viewed in this way, medical review criteria are the building blocks for science-based performance measurement and standard setting in health care.

This two-volume report describes methodologies for translating AHCPR-supported clinical practice guidelines into review criteria and performance measures, and applications of those measures in quality of care standard-setting, assessment, and improvement. The studies included in this series illustrate the challenges of and innovative approaches to implementing clinical practice guidelines, measuring guideline-related performance, and assessing the effects of putting guidelines into practice. It is our hope that these reports will not only help health care professionals integrate guidelines into their own quality improvement programs, but will also stimulate discussion about concepts and methodology that will advance the field of quality measurement.

Clifton R. Gaus, Sc.D.
Administrator
Agency for Health Care Policy and Research

Preface

Change is occurring faster than ever before, or so it seems in health care in the United States. A restructuring is under way in the financing and delivery of health care, and at the national level new roles are being debated for the Government and the private sector in this huge enterprise that constitutes one-seventh of the economy.

While there are considerable differences of opinion among the many interested parties on just how this restructuring can best be accomplished, a clear consensus is emerging on some issues related to *quality* of care, such as the following:

- Quality of care is not a given.
- There should be full public accountability for outcomes of care.
- Accountability requires measuring the quality of care provided to populations of patients.
- Public policies should promote continuous improvement of quality of care.

This document is one effort by the Agency for Health Care Policy and Research (AHCPR) in fulfilling a legislative mandate to help achieve these new approaches to quality improvement, measurement, and monitoring. In short, this document was developed to help those interested in using clinical practice guidelines—developed by AHCPR and in the private sector—to evaluate quality of care. Although many view clinical practice guidelines primarily as tools in clinical quality improvement, interest has recently been growing in the application of clinical practice guidelines to broader purposes, such as payment and reimbursement policies and legal defense. Regardless of the purpose for which clinical practice guidelines are intended to be used, it is critical that everyone who uses them understands how to do so. For *any* judgment to be made about health care delivered according to recommendations based on clinical practice guidelines, such care must be evaluated with tools that have been rigorously developed on the basis of a clearly understood methodology.

Using Clinical Practice Guidelines To Evaluate Quality of Care describes how to translate clinical practice guidelines into evaluation tools and how to use these tools to determine whether the care provided conforms to the guidelines. The methodology is not simple, yet it is nonetheless conceptually straightforward. Good measurement of care—measurement that is credible to clinicians, managers, and the public—requires meticulous attention to rigorous technique. Therefore, readers are urged not to look for shortcuts in thinking through the issues and applying the techniques.

Using Clinical Practice Guidelines To Evaluate Quality of Care has been published in two volumes. Volume 1 consists of Chapters 1–3, which provide background and overview information about clinical guidelines, their potential usefulness in quality measurement, and issues that should be considered and

understood before proceeding to develop evaluation tools. Volume 2 (Chapters 4–6) contains detailed how-to instructions for developing guideline-derived evaluation tools. They describe the state of the art in the emerging science of quality measurement. While the recommendations may seem very detailed overall, the basic steps in the processes described can be applied to beginning efforts with limited resources as well as to mature quality measurement programs with ample resources.

The appendixes provide additional useful information related to complementary efforts in outcomes measurement, validity of performance measures, statistical ratio-based measurement, and algorithm development for review criteria.

Readers are encouraged to think through the issues in deriving measurements from guidelines and to apply the methods carefully to their own guideline-related measurement efforts. Only through additional efforts will the recommendations in this volume actually improve the methodology of care assessment and thereby improve care. The Nation is unquestionably poised on the threshold of this new era of accountability in health care.

Stephen C. Schoenbaum, M.D., M.P.H.
David N. Sundwall, M.D.
Workgroup Co-chairs

Abstract

As the United States undergoes a restructuring of the way health care is financed and delivered and debates the roles of the Government and the private sector in health care, there is growing interest in examining issues surrounding quality of care—how it is defined, how it is evaluated, and how delivery of care is related to patient outcomes.

A panel of experts with a broad range of expertise and experience in measuring quality of care was convened to develop a document that would provide a single methodology for using clinical practice guidelines to evaluate the quality of health care. Volumes 1 and 2 of this document describe how to translate clinical practice guidelines into evaluation tools and how to use these tools to determine whether the care provided conforms to the recommended guidelines. A fundamental assumption of this document is that such a methodology could have broad utility for a variety of potential users, including practitioners, providers, and professional organizations.

Workgroup Members

Stephen C. Schoenbaum, M.D., M.P.H.
Co-chair
Medical Director
Harvard Community Health Plan
New England Division
Providence, RI
Internist

David N. Sundwall, M.D.
Co-chair
President
American Clinical Laboratory Association
Washington, DC
Family Physician

David Bergman, M.D.
Department of Pediatrics
Lucile Salter Packard Children's Hospital
Palo Alto, CA
Pediatrician

June M. Buckle, Sc.D., R.N.
Director of Health Information Research and Development
GMIS
Malvern, PA
Nurse Researcher

Allan Chernov, M.D.
Vice President, Medical Services
Prudential Health Care System
Southwestern Group Operations
Sugar Land, TX
Internist

Janet George, R.N., B.S.N., M.A.
Director, Quality Management
Manor HealthCare Corporation
Silver Spring, MD
Nurse

Clark Havighurst, J.D. (resigned 1993)
Wm. Neal Reynolds Professor of Law
Duke University School of Law
Durham, NC
Attorney

M. J. Jurkiewicz, M.D.
Professor of Plastic Surgery, Emeritus
Emory University
Atlanta, GA
Surgeon

John T. Kelly, M.D., Ph.D.
Chief Medical Officer and Senior Vice President of Clinical Information Services
GMIS
Malvern, PA
Physician

Sandra Metzler, M.B.A., R.R.A.
Director, Resource and Outcomes Management
Columbia/HCA Healthcare Corporation
Nashville, TN
Health Information Specialist

Christine Miaskowski, Ph.D., R.N., F.A.A.N.
Associate Professor and Interim Chair
University of California San Francisco School of Nursing
Department of Physiologic Nursing
San Francisco, CA
Nurse Researcher

Sam J. W. Romeo, M.D., M.B.A.
Senior Associate Dean
University of Southern California
School of Medicine
Los Angeles, CA
Family Physician

Paul M. Schyve, M.D.
Senior Vice President
Joint Commission on Accreditation of Healthcare Organizations
Oakbrook Terrace, IL
Psychiatrist

Bryan Simmons, M.D.
Medical Director, Quality Management
Methodist Health Systems
Memphis, TN
Hospital Epidemiologist

Patrice Spath, A.R.T., B.A.
Partner, Brown-Spath & Associates
Forest Grove, OR
Health Information Specialist

Marcia Stevic, Ph.D., R.N.
Consultant
Sun Lakes, AZ
Health Services Researcher

Constance Winslow, M.D., M.B.A.
Consultant, Health Policy and Health Management
Berkeley, CA
Internist

Steven Zatz, M.D.
Senior Vice President
US Health Care
Blue Bell, PA
Internist

Methodology Consultants

R. Heather Palmer, M.B., B.Ch., S.M.
Director, Center for Quality of Care Research and Education
Harvard School of Public Health
Boston, MA

Naomi J. Banks, M.B.A., M.Ed.
Senior Research Analyst
Center for Quality of Care Research and Education
Harvard School of Public Health
Boston, MA

Ann G. Lawthers, Sc.D.
Lecturer
Harvard School of Public Health
Boston, MA

E. John Orav, Ph.D.
Associate Professor of Biostatistics
Department of Biostatistics
Harvard School of Public Health
Boston, MA

Technical Specialists

Alice G. Gosfield, J.D.
Alice G. Gosfield & Associates, P.C.
Philadelphia, PA
Attorney

Donald M. Nielsen, M.D.
Associate Medical Director and Quality Consultant
Permanente Medical Groups
Oakland, CA
Internist/Infectious Diseases

Writer/Editor

Suzanne Wymelenberg
Cambridge, MA

AHCPR Staff

Carole Hudgings, Ph.D., F.A.A.N.
Senior Health Policy Analyst
Office of the Forum for Quality and Effectiveness in Health Care

Jill Bernstein, Ph.D.
Chief, Program Planning, Research Development, and Evaluation Branch
Office of Program Development

Linda Demlo, Ph.D.
Director, Office of Program Development

Kathleen Hastings, R.N., J.D., M.P.H.
Director, Legal Medicine Program

Kathleen A. McCormick, R.N., Ph.D.
Senior Science Advisor
Office of Science and Data Development

Randie Siegel
Managing Editor
Center for Research Dissemination and Liaison

Federal Liaisons

Galen Barbour, M.D.
Associate Chief Medical Director for Quality Management
Department of Veterans Affairs
Washington, DC

Stephen Jencks, M.D.
Clinical Advisor
Health Standards and Quality Bureau
Health Care Financing Administration
Baltimore, MD

Michael J. Kussman, M.D.
Director, Medical Quality Assurance
OASD (HA) PAQA
Washington, DC

Acknowledgments

This document is the product of the entire workgroup, which brought to the table extensive experience with measurement of health care in a variety of settings and contexts. This experience proved invaluable in the delineation of important issues to be considered in developing and applying guideline-derived measures. The efforts of the following members of the workgroup who wrote drafts of specific issues are gratefully acknowledged: Dr. June Buckle, Dr. Allan Chernov, Mr. Clark Havighurst, Dr. John Kelly, Dr. Paul Schyve, Ms. Patrice Spath, and Dr. Marcia Stevic.

Dr. Heather Palmer and her colleagues, Ms. Naomi Banks, Dr. Ann Lawthers, and Dr. E. John Orav, played a crucial role in the work of the group and in the development of this document. Dr. Palmer and her colleagues had already worked out a basic method for developing review criteria and performance measures from guidelines and had tested it in the American Medical Review Research Center project. This project involved developing medical review criteria from three of the earliest AHCPR-supported clinical practice guidelines and the DEMPAQ project (the project to Develop and Evaluate Methods for Promoting Ambulatory Care Quality), supported by the Health Care Financing Administration. Her group, with input from all workgroup members and the assistance of the workgroup's medical writer, Ms. Suzanne Wymelenberg, wrote the chapters for Volume 2 on the translation methodology. Material for Appendix A was prepared by Dr. Schyve (Appendix A, "Outcomes as Performance Measures").

Dr. Carole Hudgings from the staff of AHCPR served not only as the administrator who kept everyone on track but also as a key contributor to the substance of the work. Without her efforts and equanimity, the work could not have proceeded so smoothly, expeditiously, and productively.

Stephen C. Schoenbaum, M.D., M.P.H.
David N. Sundwall, M.D.
Workgroup Co-chairs

Contents

Executive Summary 1

1. AHCPR-Supported Guideline Development and the Activities of the Workgroup 3
Introduction 3
Background Information About AHCPR 3
The AHCPR-Supported Clinical Practice Guidelines 4
Selection of Clinical Practice Guideline Topics 4
Methodology for Developing AHCPR-Supported Guidelines 4
Guideline Dissemination 6
Other Sources of Guidelines 6
The AHCPR's Workgroup 7
References 8

2. Guideline-Derived Quality Evaluation 9
Introduction 9
Basics of Quality Evaluation 9
Improving the Quality of Health Care 10
A Context for Understanding the Workgroup's Proposed Methodology 12
Reference 13

3. Issues To Consider Before Developing and Implementing Evaluation Tools 15
Introduction 15
General Issues 15
Multiple Users and Uses 15
Who Should Participate 16
Variation in Settings 16
Variation in Resources 17
Variation in Data and Documentation 17
Training and Education 18
Variation in Process 19
Issues Related to the Translation Methodology 19
Need To Be Selective 19
Evaluating the Guideline Itself 21

Reviewing Periodically 21
Planning for Implementation 21
Potential Barriers 22
Integration With Ongoing Quality Management Activities 22
Possible Impact on Health Care Practitioners 23
A Tool, Not a Rule 23
Standards of Quality 23
Process Versus Outcome Measurement 23
Public Disclosure 24
Legal Issues 24
Summary 25
References 25

Appendix
A. Outcomes as Performance Measures 27

Glossary 35

Peer Reviewers 39

Tables
1.1. Released AHCPR-supported clinical practice guidelines 5
1.2. Basic steps in developing AHCPR-supported clinical practice guidelines 6
2.1. Definitions 11
2.2. Example of a clinical practice guideline-derived evaluation tool for determining quality of care for postoperative pain control 13
3.1. Potential types and sources of data 17
3.2. Attributes of good practice guidelines 20

Figures
2.1. A model of care as a basis for quality evaluation 10
2.2. The quality management cycle 11
2.3. Relationship between clinical practice guidelines and evaluation tools 12

Executive Summary

As the United States restructures its health care financing and delivery systems and debates the roles of Government and the private sector in the delivery of health care, a consensus is emerging on several issues:

- Quality of care is not a given.
- There should be full public accountability for outcomes of care.
- Accountability requires measurement of the quality of care provided to populations of patients.
- Public policies should promote continuous improvement of quality of care.

To fulfill its legislative mandate to help achieve improvements in health care delivery, measurement, and monitoring, AHCPR convened a multidisciplinary expert panel to examine ways in which clinical practice guidelines can be used to evaluate quality of care. In order for any judgment to be made about health care delivered according to recommendations based on clinical practice guidelines, such care must be evaluated with tools that have been developed on the basis of a clearly understood methodology. This document considers the issues involved in developing and using evaluation tools. A second volume provides the methodology.

This project was guided by extensive discussion and review of materials by the workgroup; by an open meeting to solicit comments and information from professional health care organizations, specialty societies, and other interested individuals; and by extensive peer review by individuals and organizations.

Volume 1 (Chapters 1–3) presents a discussion of guideline-derived quality evaluation in which the basic tools for improving health care—medical review criteria, performance measures, and standards of quality—are briefly defined and described. Discussion follows on a number of issues that users should consider before developing and implementing evaluation tools based on clinical practice guidelines. Among these issues are several general questions, such as the potential uses and users for this evaluation methodology; the nature and training of the multidisciplinary group that will apply the methodology to a particular health care issue; and the variations in settings, resources, and data and documentation that may affect the evaluation process. A second broad category of issues that users should consider before developing evaluation tools has to do with the translation methodology itself. These issues relate to selecting a topic for evaluation; developing an implementation plan for the evaluation tools; understanding that developing and using evaluation tools is a dynamic process; and identifying potential barriers to implementing evaluation tools. The third category of issues to consider comprises the uses to which the information derived from the application of evaluation tools will be put and the impact of that information on health care practitioners themselves.

Volume 1 concludes with an appendix containing supplementary information on the use of outcomes as performance measures.

The heart of *Using Clinical Practice Guidelines To Evaluate Quality of Care* lies in Volume 2, in which Chapters 4 and 5 lay out an overview and explain step-by-step the methodology for translating clinical practice guidelines into evaluation tools. Volume 2 includes a checklist (Chapter 6) that

provides performance review committees with a means for checking the completeness of their work and documenting their decisions. Also included are appendixes on the development of a process to determine the validity of a performance measure, on selected statistical issues, and on methods for constructing algorithm flowcharts.

1. AHCPR-Supported Guideline Development and the Activities of the Workgroup

Introduction

This chapter presents background information on AHCPR and the development of clinical practice guidelines. AHCPR-supported clinical practice guidelines provide an important and useful starting point for generating medical review criteria, performance measures, and standards of quality, which collectively are referred to as *evaluation tools* that can be used to assess quality of care. Also described in this chapter are the formation, composition, and processes followed by the AHCPR-supported Workgroup To Develop Methods for Deriving Medical Review Criteria, Standards of Quality, and Performance Measures From Guidelines. The activities of the workgroup led to this document.

Background Information About AHCPR

AHCPR was established by the Omnibus Budget Reconciliation Act (OBRA) of 1989 (Public Law 101-239) as the successor to the National Center for Health Services Research and Health Care Technology Assessment. As the newest component of the Public Health Service, AHCPR was created in part to respond to (1) concerns about the variability in health care practices and (2) uncertainty about the effectiveness of health care services. The Agency's mission is to improve the quality, appropriateness, and effectiveness of health care and improve access to health services.

AHCPR carries out this mission by supporting and conducting scientific research, facilitating the development of clinical practice guidelines, and assessing health care technologies. Agency-funded research explores the availability, quality, and costs of health care services and ways to improve the effectiveness and appropriateness of clinical practice, including prevention of disease. In-house researchers analyze a range of health care policy and cost issues and evaluate the risks and effects of specific health care technologies.

The Agency's Medical Treatment Effectiveness Program (MEDTEP) addresses questions of effectiveness and appropriateness of clinical practice. Components of MEDTEP include patient outcomes research, clinical practice guidelines, health care data bases for research, and dissemination of research and guideline information.

The Office of the Forum for Quality and Effectiveness in Health Care (the Forum) is one organizational component of AHCPR. In support of the overall Agency mission, the Forum is directed by the 1989 OBRA legislation to arrange for the development, periodic review, and updating of the following: (1) clinically relevant guidelines that may be used by physicians, educators, and health care practitioners to assist in determining how diseases, disorders, and other health conditions can most effectively and appropriately be prevented, diagnosed, treated, and managed clinically; and (2) standards of quality, performance measures, and medical review criteria through which health care providers and other appropriate entities may assess or review the provision of health care and ensure its quality.

The AHCPR-Supported Clinical Practice Guidelines

Soon after its creation, AHCPR sought advice from the Institute of Medicine (IOM) regarding the development of clinical practice guidelines. The IOM convened a special committee that defined key terms, identified desirable attributes of good guidelines, and provided technical assistance for other aspects of guideline development (IOM, 1990).

As defined by IOM, practice guidelines are "systematically developed statements to assist practitioner and patient decisions about appropriate health care for specific clinical circumstances" (IOM, 1990, p. 38). This definition highlights several distinctive aspects of the AHCPR-supported clinical practice guidelines. The guidelines represent "systematically developed statements," that is, statements generated from use of a systematic, rigorous, and explicit methodology, including a review of all available scientific evidence for the topic under study. The AHCPR-supported guidelines are intended to assist both practitioners and patients in making health care decisions. In addition, the guidelines focus on clinical conditions rather than on technical procedures. Within this broad definition, the purpose of the AHCPR-supported clinical practice guidelines is to enhance the quality, appropriateness, and effectiveness of health care, using the best available scientific evidence and expert opinion.

Selection of Clinical Practice Guideline Topics

In selecting topics for guidelines to be developed, AHCPR considers many factors, including the following:

- Potential for reducing clinically significant and unexplained variations in services and procedures used in the prevention, diagnosis, treatment, management, or outcomes related to the clinical condition.
- Number of individuals affected by the condition.
- Adequacy of scientific evidence on which to develop a guideline.
- Amenability of a particular condition to prevention.
- Specific needs of the Medicare and Medicaid populations.
- Cost of the condition to all payers, including patients.

Potential guideline topics are suggested by a variety of health care experts and professional organizations. In 1993, AHCPR published a *Federal Register* notice in which it invited comments on guideline topics under consideration and requested suggestions for additional topics. Expert advice is also sought on prioritizing guideline topics. The topics selected by AHCPR for guideline development cover a broad range of clinical problems and relate to the work of many different types of health care professionals. Guidelines released to date are listed in Table 1.1.

Methodology for Developing AHCPR-Supported Guidelines

The AHCPR-supported clinical practice guidelines are developed by private sector expert panels who use a rigorous and explicit methodology. Members of the multidisciplinary guideline panels are selected from nominations received in response to a *Federal Register* notice announcing the formation of a panel on a specific topic. Panel chairs and members are selected according to published criteria, including expertise in the clinical condition under study. Each panel consists of approximately 15 members who represent a broad range of professional disciplines and clinical experience relevant to the clinical condition under study and a diversity of geographic locations and ethnic origins. All guideline development panels include health care consumers.

The AHCPR's guideline methodology used by the expert panels was formulated with the assistance of the IOM and individual experts in clinical

Table 1.1. Released AHCPR-supported clinical practice guidelines

Guideline number and title		Release date
1.	*Acute Pain Management: Operative or Medical Procedures and Trauma*	March 5, 1992
2.	*Urinary Incontinence in Adults*	March 23, 1992
3.	*Pressure Ulcers in Adults: Prediction and Prevention*	May 18, 1992
4.	*Cataract in Adults: Management of Functional Impairment*	February 25, 1993
5.	*Depression in Primary Care: Volume 1. Detection and Diagnosis. Volume 2. Treatment of Major Depression*	April 14, 1993
6.	*Sickle Cell Disease: Screening, Diagnosis, Management, and Counseling in Newborns and Infants*	April 27, 1993
7.	*Evaluation and Management of Early HIV Infection*	January 20, 1994
8.	*Benign Prostatic Hyperplasia: Diagnosis and Treatment*	February 8, 1994
9.	*Management of Cancer Pain*	March 2, 1994
10.	*Unstable Angina: Diagnosis and Management*	March 15, 1994
11.	*Heart Failure: Evaluation and Care of Patients With Left Ventricular Systolic Dysfunction*	June 29, 1994
12.	*Otitis Media With Effusion in Young Children*	July 14, 1994
13.	*Quality Determinants of Mammography*	October 19, 1994
14.	*Acute Low Back Problems in Adults*	December 8, 1994
15.	*Treatment of Pressure Ulcers*	December 21, 1994

practice guideline development, analysis of scientific research, information analysis, and evaluation research. This methodology requires that all scientific evidence related to the clinical condition under study be considered, that benefits and harms of various diagnostic modalities and interventions be identified, that the consequences of different treatment options be weighed, and that the scientific evidence and expert judgment supporting the guideline statements be described and documented explicitly.

This methodology emphasizes the importance of comprehensively evaluating the empirical evidence demonstrating effectiveness, significant outcomes (including those important to patients), benefits, and risks. Through their analysis of the scientific literature, panels seek to identify which health care interventions for a particular clinical condition result in what outcomes (both beneficial and harmful). The expert judgment of the panel is used as the basis for guideline recommendations only in the absence of scientific evidence, or when the scientific evidence is weak or conflicting (Woolf, 1991). In developing the AHCPR-supported guidelines, expert panels go through a series of basic steps (Table 1.2).

AHCPR-supported guideline panels focus on the health interventions and the health outcomes that are associated with the clinical condition of interest. Health interventions are "any action taken to modify the health of an individual or collection of individuals" (Eddy, 1991, p. 28). Health outcomes are defined as "outcomes that people can experience (feel physically or mentally) and care about. They relate to the length and quality of life, including death, functional disability, appearance, pain, anxiety, reassurance, and peace of mind" (Eddy, 1991, p. 54). For additional information about outcomes and the relationship between outcomes and interventions or processes of care, see Appendix A.

An important characteristic of the AHCPR-supported clinical practice guidelines is that they address patient preferences. The guidelines specify certain decision points at which the benefits and risks of interventions are presented to the patient

Table 1.2. Basic steps in developing AHCPR-supported clinical practice guidelines

- Define the scope of the guideline.
- Define the major questions regarding the clinical condition that the guideline will address.
- Review and analyze the available scientific research for each question.
- Review estimates of important patient outcomes that will be influenced by the interventions for the clinical condition.
- Review benefits and harms of each intervention.
- Review current and potential health care costs associated with the guideline. Where cost information is available and reliable, provide costs of alternative strategies for the prevention, diagnosis, treatment, and management of the condition.
- Invite information and comments on the guideline topic from professional and consumer organizations, researchers, and others at a public open meeting.
- Prepare the guideline draft on the basis of available empirical evidence and professional judgment when empirical evidence is insufficient.
- Submit the draft guideline for independent review by outside experts.
- Revise the draft guideline on the basis of analysis of comments and information received.
- Prepare the guideline in several formats appropriate for use by providers, consumers, educators, researchers, and others.
- Periodically review the new scientific evidence related to the clinical condition, and update the guideline when such evidence indicates a need.

so that the patient's preferences can be taken into account in the decisionmaking process.

Guideline Dissemination

Once completed, the AHCPR-supported clinical practice guidelines are widely disseminated to a variety of audiences. The AHCPR-supported guidelines comprise several documents, each created to meet the needs of a particular audience. The guideline documents are as follows:

- **Clinical practice guideline.** This version contains the specific guideline statements and recommendations, algorithms, summary of evidence tables, and pertinent references. It is intended for use by health care practitioners as a reference for clinical decisionmaking.

- **Quick reference guide for clinicians.** This document, an abbreviated version of the clinical practice guideline, includes summary points of prevention, diagnosis, and treatment. It is for day-to-day use by health care practitioners.

- **Consumer version** (in English and Spanish). This version for patients, caregivers, and parents of young patients describes the condition, presents treatment options with benefits and risks in easy-to-understand terms, and suggests questions to ask health care practitioners. The information is intended to help patients make decisions about treatment alternatives.

- **Guideline technical report.** This technical version contains all background and supporting materials for the clinical practice guideline, including algorithms, methodology, literature review, summary of scientific evidence tables, references, and a comprehensive bibliography. It is for health care practitioners, researchers, educators, professional organizations, and others.

Other Sources of Guidelines

In addition to the AHCPR-supported clinical practice guidelines, there are other sources of guidelines. Among the most readily available sources are the practice guidelines and practice parameters developed by physician and nurse organizations and by other health care organizations. For example, approximately 1,800 practice parameters developed by almost 75 national physician specialty organizations and other groups are available (AMA, 1995). Other guidelines have been developed by the U.S. Preventive Services Task Force (1989). While the focus of this document is the use of the AHCPR-supported clinical practice guidelines as a basis for deriving medical review criteria, performance

measures, and standards of quality, the workgroup believes that the processes and methodology described could be used to derive these evaluation tools from guidelines developed by other organizations.

The AHCPR's Workgroup

The second legislative mandate for the Forum is the development of medical review criteria, performance measures, and standards of quality. One fundamental strategy used by the Forum to address this mandate was to convene the multidisciplinary workgroup that developed this document. Early discussions among the AHCPR staff and representatives from other Federal agencies resulted in the recognition that no documented methodologies existed for translating clinical practice guidelines into evaluation tools. Also acknowledged was that such methodologies, when developed and documented, could be helpful to a variety of possible users, including practitioners, providers, professional organizations, and others. An essential assumption underlying the workgroup's formation and charge by the Forum was that translation methodologies could be devised to meet the needs of various potential users for evaluating quality of care. The workgroup subsequently concluded that a single translation methodology could be devised to accommodate the needs of multiple users.

Establishment of the workgroup was announced in a *Federal Register* notice dated January 23, 1992, which called for nominations of experts to serve as workgroup members. More than 50 nominations of qualified individuals were received from health care organizations, professional societies, and experts in health care quality. Two co-chairs and 16 members were selected for the workgroup after an extensive review of the nominations. These 18 men and women have a broad range of expertise and experience in measuring quality of care and represent diverse clinical practice settings, research expertise, education, and geographic locations as well as diverse disciplines and specialties. In addition, the Agency contracted with Dr. R. Heather Palmer to be the methodologic consultant to the workgroup.

The workgroup convened its first meeting in March 1992 in Charleston, SC. Thereafter, five meetings were held in the Washington, DC, area between June 1992 and September 1993. The AHCPR charged the workgroup as follows:

> The Workgroup is charged with identifying issues associated with and methods for translating clinical practice guidelines into medical review criteria, standards of quality, and performance measures. This work is intended to be useful to a variety of organizations engaged in developing guidelines, evaluating health care, or improving the quality of health care.

In the course of its meetings, the workgroup proceeded to meet its two charges—to develop *methods* for translating clinical practice guidelines into evaluation tools, and to consider and describe *issues* associated with developing and using evaluation tools. Dr. Palmer began by outlining a specific translation methodology that built on work that had begun as part of an AHCPR contract to the American Medical Review Research Center to develop medical review criteria for three specific AHCPR-supported guidelines. Dr. Palmer's initial proposal of a framework and all subsequent iterations of the methodology were reviewed in depth by the workgroup members, who contributed significantly to its refinement for general usage.

Concurrently, the workgroup identified, discussed, and documented the issues relating to an organization's development and use of guideline-derived evaluation tools. The initial list of issues was derived from workgroup discussion; this list was expanded and refined during subsequent meetings and reviews of draft chapters. Early drafts regarding specific issues were prepared by several workgroup members. These drafts provided the basis for subsequent discussion, and much of this information was incorporated into Chapter 3.

The workgroup's open meeting, held September 1, 1992, provided important information that

proved helpful for framing, discussing, and describing relevant issues. This open meeting, announced in a Federal Register notice, was an opportunity for the workgroup to receive comments and information from the public about its activities. Attendance at this open meeting reflected the great interest of the health care community in the development and use of evaluation tools. A total of 22 professional health care organizations, specialty societies, and interested individuals presented oral or written testimony for the workgroup's consideration. Using the nominal group technique, the workgroup members identified 52 major themes from all testimony. These themes were considered as the methodology was refined, and the related issues were explored and described.

In July 1993, AHCPR invited 111 persons and organizations to review a draft of this document. Included were individuals who had been nominated for workgroup membership, persons and organizations who had presented to the workgroup at the open meeting, and other persons known to AHCPR or to workgroup members to have interest and expertise in quality measurement in a variety of settings. Ninety-two of these persons indicated an interest in commenting on the draft document, and 61 of those returned comments for consideration by the workgroup. Discussion of these comments was the substance of the workgroup's last meeting in September 1993 and the basis for the final document.

References

American Medical Association (AMA). Directory of practice parameters: titles, sources, and updates. Toepp MC, Kuznets N, Herrera S, editors. Chicago: American Medical Association; 1995.

Eddy DM. A manual for assessment of health practices and designing practice policies: the explicit approach. Philadelphia, PA: The American College of Physicians; 1991.

Institute of Medicine (IOM), Committee on Clinical Practice Guidelines. Clinical practice guidelines: directions for a new program. Field MJ, Lohr KN, editors. Washington, DC: National Academy Press; 1990.

U.S. Preventive Services Task Force. Guide to clinical preventive services: an assessment of the effectiveness of 169 interventions. Baltimore, MD: Williams & Wilkins; 1989.

Woolf SH. AHCPR interim manual for clinical practice guideline development. AHCPR Pub. No. 91–0018. Rockville, MD: Agency for Health Care Policy and Research, Public Health Service, U.S. Department of Health and Human Services; May 1991.

2. Guideline-Derived Quality Evaluation

Introduction

Over the past half-century or more, medical progress has consisted primarily of the development of new technical information about disease processes, new diagnostic techniques, and new treatments. Relatively less attention has been paid to developing systematic information about how health care is actually delivered to individual patients or populations of patients. More recently, as this type of information has been sought, variation in clinical practices has been documented and clearly demonstrates that persons with the same basic condition often receive different evaluations or treatments. These differences in the processes of health care lead to variation in the costs of care and, more important, may lead to differences in the outcomes of care for the patients being treated.

To improve significantly the delivery of health care and achieve optimal outcomes, more information about care is needed. For example, clinicians and patients need information on the outcomes of care for large groups of patients to help them make decisions about individual patient care; clinicians and provider organizations need information about individual and comparative performance to assess and improve care; and payers and purchasers need information to help them decide what to pay for and whether they are receiving the high-quality care they are paying for.

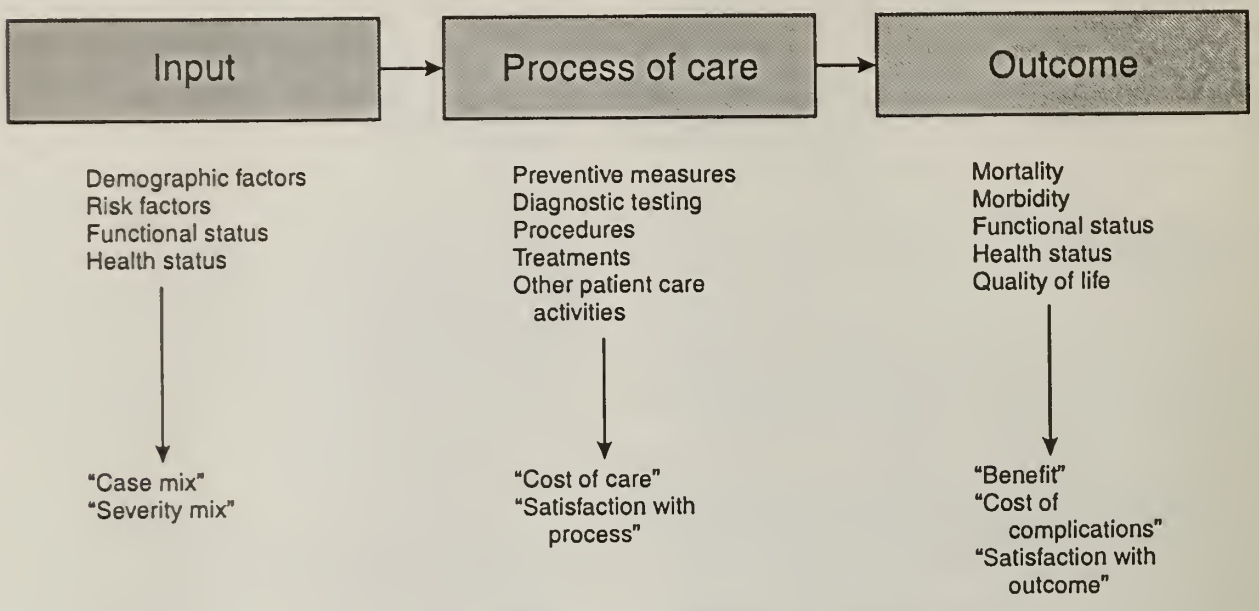
Considerations such as these lay behind Congress's establishment of the AHCPR in OBRA 1989. These considerations also led to the Agency's mandate to study outcomes of care and to facilitate development of clinical practice guidelines, medical review criteria, performance measures, and standards of quality (see Chapter 1). The development of clinical practice guidelines and guideline-derived measures of clinical performance, which is by no means limited to the efforts of AHCPR, is an important step in the transformation of health care delivery into an industry based on facts about the relationship between processes of care and outcomes of care.

Basics of Quality Evaluation

To understand the workgroup's approach to delineating issues and methods associated with translation of clinical practice guidelines into evaluation tools, it is desirable to have a framework or model for understanding health care in measurable terms (Figure 2.1). The objective of health care is to achieve good or improved outcomes for patients. Quality measurement should determine whether the processes of care provided to a single patient or population of patients (1) achieved good outcomes (outcomes measurement) or (2) represented those processes that are thought or known to be associated with achievement of good outcomes (process measurement).

Processes of care include elements such as preventive measures, diagnostic tests, treatments, and other patient care activities. In the aggregate, these processes account for most of the "cost of care" (see Figure 2.1) The way in which these processes are delivered leads to an important component of patients' satisfaction or dissatisfaction with care. Outcomes typically include morbidity, mortality, health status, functional status, and quality of life. These outcomes can be aggregated into the "benefits" of care or the "cost of complications." The nature of these outcomes can also influence the patient's overall satisfaction or dissatisfaction with

Figure 2.1. A model of care as a basis for quality evaluation



care. To make determinations about processes or outcomes, one must take into account the individual and collective characteristics of the patients who received care—the inputs to care, such as demographic or risk factors, and the patients' baseline functional status and health status. These input factors determine the "case mix" or "severity mix" of the population that is receiving care. Although doing so is no trivial task, each of these elements of the health care model can be quantified.

For the most part, practice guidelines specify processes of care that are known or believed to be associated with good outcomes. Accordingly, most measurement derived from clinical practice guidelines is process measurement. A guideline-derived performance measure is likely to be a valid quality indicator if these three conditions are met: (1) the relationship between process and outcome is known (for further discussion, see Appendix A); (2) that relationship formed the basis for the practice guideline; and (3) the review criteria have been derived "faithfully" from the guideline. The last of these conditions, faithful derivation of the review criteria from guidelines, is most likely to occur if the issues important in developing review criteria have been considered and an explicit method has been followed.

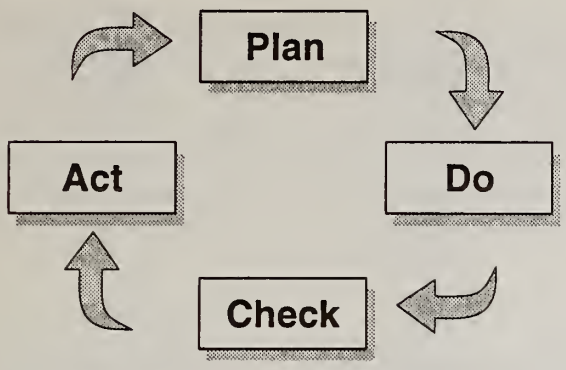
Improving the Quality of Health Care

In assessing and monitoring the quality of care, the health care industry is moving from review of individual cases to a broader, population-based perspective. Peer review organizations (PROs) and accrediting organizations are moving toward population-based assessments of quality and quality improvement. In addition, an important component of all health care reform proposals is the provision of care to populations, not just individuals.

Quality has become a critical issue in almost all sectors of business, industry, and organizations. It is now recognized that error, waste, and rework lead to higher costs and less customer satisfaction. Hence, there is now an interest in formally adopting quality improvement or management techniques in industry in general and in the health care sector in particular (Berwick, 1989).

At the heart of quality improvement methodology is the so-called quality management cycle

Figure 2.2. The quality management cycle



(Figure 2.2), or Shewhart cycle. In its simplest form, this cycle consists of *planning* a new or improved process that is expected to yield an improved outcome, *doing* or implementing the process, *checking* or measuring to see whether the process is working as planned and is having the expected outcomes, and then *acting* on the information that has been obtained to reassess the plan or continue implementing it.

Clinical practice guidelines, medical review criteria, performance measures, and standards of quality are tools that can be used to improve health care (see Table 2.1 for the workgroup's definitions of these terms). Clinical practice guidelines compile available scientific evidence and expert opinion into specifications for the processes of care that are likely to maximize effectiveness (good outcomes).

Development of clinical practice guidelines can be considered the planning phase of improving care, or the first step of the quality management cycle. It is usually performed by a central agency or organization, whereas implementation of clinical practice guidelines, the second step of the quality management cycle, is a task for clinicians and others working directly in health care delivery systems.

Medical review criteria and performance measures are used to determine the degree to which care has followed specified processes and whether the desired outcomes have been achieved. Measuring performance is the way to determine that care conforms with the recommendations in clinical practice guidelines (i.e., that appropriate processes of care have been implemented). When the measurements are compared with standards of quality, one can determine whether the care met predetermined boundary conditions or targets, information that helps in deciding what actions, if any, should be taken next. Measurement (the third step of the quality management cycle) and action (the fourth step) are tasks for any party interested in quality of care and its improvement (e.g., clinicians, health care delivery systems, payers, regulators, and accreditors).

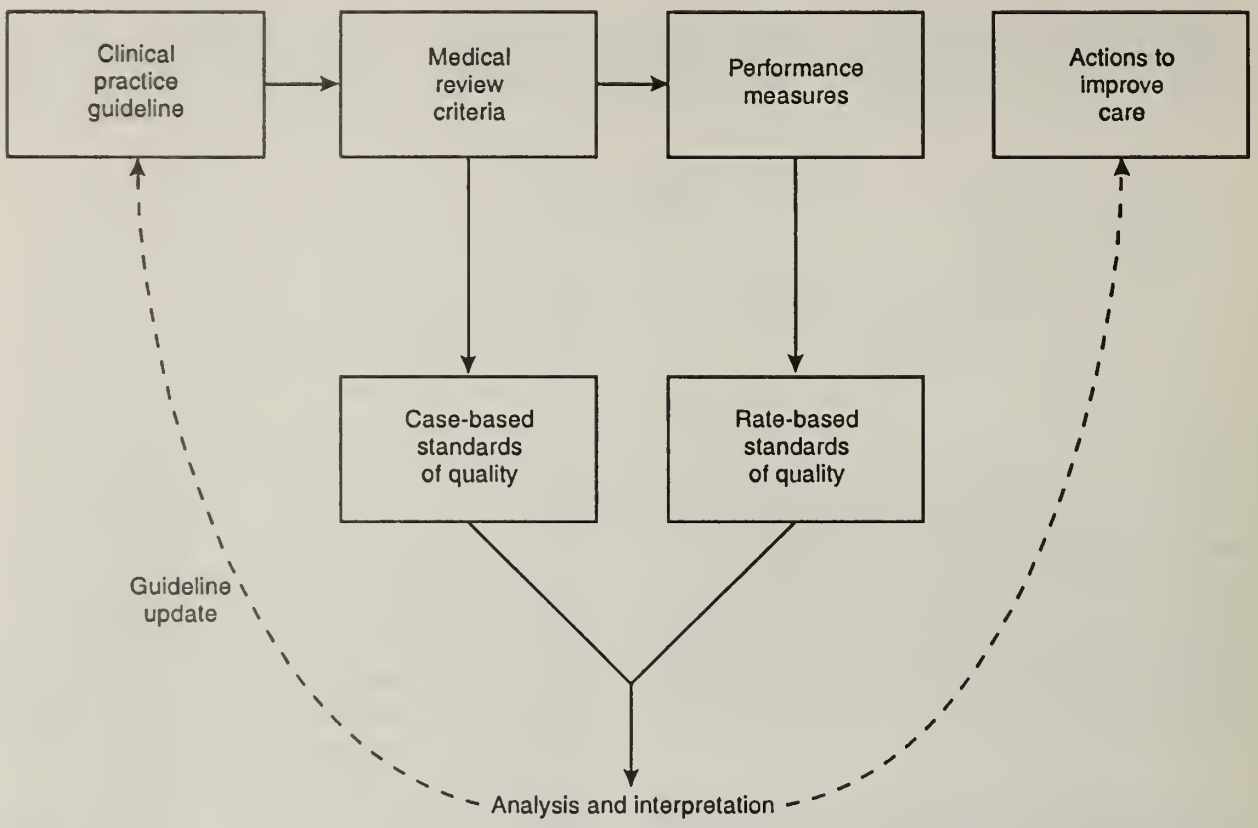
The specific interrelationships among clinical practice guidelines, medical review criteria, performance measures, and standards of quality are shown in Figure 2.3 and explained further in Chapter 4. An example of how the elements might be applied is shown in Table 2.2.

Table 2.1. Definitions

Clinical practice guidelines	Systematically developed statements to assist practitioners' and patients' decisions about health care to be provided for specific clinical circumstances.
Medical review criteria	Systematically developed statements that can be used to assess specific health care decisions, services, and outcomes.
Performance measures	Methods or instruments to estimate or monitor the extent to which the actions of a health care practitioner or provider conform to the clinical practice guideline.
Standards of quality	Authoritative statements of (1) minimum levels of acceptable performance or results, (2) excellent levels of performance or results, or (3) the range of acceptable performance or results.

SOURCE: Adapted from Institute of Medicine, Committee on Clinical Practice Guidelines. Clinical practice guidelines: directions for a new program. Field MJ, Lohr KN, editors. Washington, DC: National Academy Press, 1990.

Figure 2.3. Relationship between clinical practice guidelines and evaluation tools



A Context for Understanding the Workgroup's Proposed Methodology

This document should be helpful to a variety of users who have an interest in clinical practice guidelines, evaluation, and improvement of care. Nevertheless, each user should be aware that the intention of the workgroup is to promulgate a general methodology for getting from practice guidelines to quality evaluation, and not to instruct users on special or particular applications.

The following examples illustrate what the workgroup has tried to achieve as well as the limitations of its work. In the future, a multihospital system may want to apply guideline-derived standards of quality to the care provided by its clinicians. The workgroup's methodology should help the hospital system translate the practice guideline into medical review criteria and performance measures. The workgroup has also outlined the issues the hospital system must face in doing so and in developing standards of quality. The hospital system, however, must decide just which review criteria, performance measures, or standards of quality should be developed; how to determine the specific level of performance that is to become the standard of quality; and what action to take to improve clinical care (e.g., adopt new processes for care or institute training to help personnel adhere better to existing processes). Similarly, a nursing home may want to evaluate and improve its care in accordance with practice guidelines. The workgroup has outlined the types of staff skills that are likely to be necessary to do this and the attributes of practice guidelines that are likely to be associated with a successful evaluation. After considering the issues in selection of guidelines, the nursing home's management can decide which guidelines to work with. Then, the workgroup's methodology should be use-

Table 2.2. Example of a clinical practice guideline-derived evaluation tool for determining quality of care for postoperative pain control

Clinical practice guideline recommendation	Pain should be assessed and documented routinely at regular intervals postoperatively, as determined by the operation and the severity of pain (e.g., every 2 hours while awake for 24 hours after surgery).
Medical review criterion	For the patient recovering from surgery, the patient's pain was assessed and documented every 2 hours while awake for the first 24 hours following surgery.
Performance measure	Calculate the following for consecutive surgical patients seen over a 6-month period: the number of patients whose pain was assessed and documented every 2 hours while awake. The performance measure is: $\frac{\text{\# of cases with criterion met}}{\text{\# surgery cases}} \times 100$ percent
Standard of quality	A performance rate of 95 percent or less triggers a review to determine how to improve assessing and documenting the patient's pain status every 2 hours while awake for the first 24 hours postoperative hours.

ful for deriving performance measurements from the guidelines. Ultimately, the nursing home management will be responsible for determining whether the care needs improvement.

The workgroup understands that measuring quality of care can be very difficult. Data collection systems are often inadequate or inefficient. The workgroup has tried to be realistic about currently existing systems and has recommended methods that can be used in today's environment. At the same time, the workgroup fully expects that as knowledge and scientific evidence increasingly become the basis of health care, new systems will be developed to facilitate the gathering of more and better information about care processes and their outcomes; for example, there will be increasing computerization of data and richer sources of data. Accordingly, the methodology outlined in this document anticipates the development of better technology to support quality measurement and is not confined merely to the most commonly available information. Nevertheless, using the workgroup's methodology and existing systems should enable a considerable amount of care to be evaluated in relation to existing clinical guidelines.

Reference

Berwick DM. Continuous improvement as an ideal in health care. N Engl J Med 1989;320:53–6.

3. Issues To Consider Before Developing and Implementing Evaluation Tools

Introduction

This chapter identifies and discusses the major issues related to translating clinical practice guidelines into evaluation tools and the possible uses of these tools to evaluate quality of health care. Three types of issues are discussed: general issues, issues related to the translation methodology, and possible impact on health care practitioners. Although the workgroup identified many of these issues, most were drawn from the public testimony and the written comments provided by a large number of interested observers and organizations. *It is important for readers to consider these issues while reading the translation methodology (Volume 2, Chapters 4 and 5) and throughout the process of developing and using guideline-derived evaluation tools.*

General Issues

Multiple Users and Uses

The focus of this document is on issues and methods for developing and implementing guideline-derived evaluation tools for the purpose of **evaluating and improving quality of care**.

Many individuals and organizations are interested in evaluating care for this purpose, including the following:

- **Individual physicians, nurses, and other health care professionals.**
- **Provider organizations**, such as hospitals, managed care organizations, ambulatory practices, and nursing homes.
- **Purchasers** of health care, both public and private.
- **Consumers** of health care, who are taking a more active and visible role in evaluating the quality of health care.
- **Regulators** of health care, both State and Federal; regulatory reviews currently focus more on performance related to quality of care than on organizational and financial issues.
- **Accrediting and certifying agencies.**
- **Managers** of health care, especially those involved in managed care organizations.
- **Researchers** who wish to define and examine attributes of quality of care.
- **Health policy analysts** and others who set or evaluate health policy.
- **Other interested parties**, including attorneys, economists, journalists, and politicians, who do not provide, purchase, or regulate care but have a powerful stake in measuring performance and evaluating the delivery of health care services.

Although each of these groups can use clinical practice guidelines to derive quality-of-care evaluation tools in order to improve care, each user's specific interests may lead to differences in the scope, focus, or content of the instruments. That is, it should not be assumed that a guideline will yield a single unique set of medical review criteria that

will meet the needs of all users who have the common purpose of improving quality. *For each set of criteria to be valid, however, it must be derived by using a rigorous methodology as described herein.*

This document is designed to assist all users who wish to evaluate care in order to improve it. However, the methods and issues described may also be useful to users who develop instruments to evaluate the quality of care for purposes other than improving quality. The possible uses of guideline-derived evaluation tools for other purposes (e.g., in malpractice liability) raised a series of concerns for those who testified at the public meeting. The use of these evaluation tools for other purposes would be facilitated, and the process less fragmented, if all users adopted the same framework for developing such instruments and used the terminology presented in this document.

The workgroup believes that the methods and issues explored in this document are adaptable to guidelines developed by organizations other than AHCPR. Evaluation tools could be developed from any clinical practice guideline if the following three conditions are met:

1. The guideline is developed by using available scientific evidence and an explicit, rigorous methodology.

2. The user understands the process by which it was developed.

3. The user is confident that the underlying scientific evidence of each guideline statement provides linkages between interventions and health outcomes.

Who Should Participate

Developing and implementing guideline-derived evaluation tools requires expertise from several disciplines. Ideally, a defined, multidisciplinary group would commit to the project over a period of time. However, the number and types of individuals involved are not as important as the individuals' expertise, skills, experience, and interest. The goals of this group may vary depending on the type of organization developing the evaluation tools and the specific purpose for which the evaluation tools are to be applied. Because of the diversity in the purpose and use of evaluation tools, the types of experts needed for development and implementation also vary.

The workgroup recommends that a multidisciplinary team develop and implement the evaluation tools. At a minimum, this team should have skills in clinical practice, quality review and measurement, and health information management. The team should include health care clinicians (such as physicians, nurses, and others whose performance will be measured) as well as quality management professionals and health information management personnel. When appropriate and feasible, statisticians and health services researchers should be included. Skills and experience in quality evaluation are likely to be more important than expertise in specific disciplines. For "buy-in," it is considered essential that clinicians whose care is being evaluated be represented and be involved in developing and implementing the evaluation tools.

If at all possible, input should be sought from the developers of the guideline from which the evaluation tools are being derived. Guideline developers, such as members of the AHCPR-supported guideline panels, are valuable in the process of developing guideline-derived evaluation tools. Because guideline developers are familiar with the content and intent of the guideline, they can provide useful information about the most important guideline recommendations, explain the intent of specific guideline recommendations, determine whether the integrity of the guideline is preserved in the review criteria (see the section of this chapter titled "Evaluating the Guideline Itself"), and ascertain whether the evaluation tools actually measure the care processes recommended by the guideline.

Variation in Settings

In general, the AHCPR-supported clinical practice guidelines are not specific to any one patient care

setting; rather, they relate to the care of patients with a specific clinical condition, regardless of setting. For example, the guideline for urinary incontinence in adults (Diokno, McCormick, Colling et al., 1992) may apply to patients in ambulatory, long-term care, hospital, or rehabilitation settings. Evaluation tools derived from an AHCPR-supported guideline may thus be developed for care delivered in similarly diverse settings. The *process* of developing and implementing evaluation tools may differ somewhat among settings, depending on the resources available. However, the *methodological steps* outlined in Chapters 4 and 5 apply to all settings, whether the organization wishes to conduct a review of patient charts or use a computer-assisted analysis of multiple records.

Variation in Resources

There is also considerable variation in the resources (human, financial, data, information technology) available to users to develop and implement evaluation tools. For example, a small or rural hospital may have limited personnel available to design and implement quality evaluation. On the other hand, a large, multihospital system may be able to assemble a multidisciplinary team consisting of physicians, nurses, and experts in quality and health information management. The multihospital system may also have access to computer-assisted analyses of large data sets. If human resources are not readily available, organizations should at a minimum obtain the necessary special expertise on a consultation basis. Some organizations (e.g., hospitals, group practices) may also find it valuable to join with other sites in a consortium to maximize the use of resources, increase the sample size of cases undergoing review, and obtain access to sufficient data for comparative analysis and study. For discussion of the costs of review see the section in Chapter 4 titled "Costs of Review."

Variation in Data and Documentation

Different users will have or seek access to different types of data, depending on the specific purpose for the evaluation tools. Categories of data useful for measuring quality include administrative (claims) data, patient health records, and patient surveys. Some of these data sources may not be readily available in some settings. In addition, certain data sources are better than others for certain data types. For example, data describing patient preferences may be best collected by means of patient surveys. Individual health care practitioners and health care institutions typically have access to the health records of patients they manage, and to the patients themselves. Organizations such as managed care firms and insurance companies may have access to patient health records from a large number of practitioners and institutions. These organizations may also have access to claims files or other administrative computerized data bases containing information abstracted from health records. It is essential to match the data needed to measure performance with the appropriate sources of such data. See Table 3.1 for examples.

Measurement based on an organization's data depends on the adequacy of the information available

Table 3.1. Potential types and sources of data

Sample data type	Potential data source				
	Administrative (claims) data	Hospital patient record	Office/agency patient records	Patient survey	Institution policies and procedures
Patient preferences				X	
Surgical procedure performed (e.g., cataract extraction)	X	X			
Medication administration (e.g., morphine)		X	X		X

in the data sources. As guideline-derived quality measurement becomes more common, it is critical that documentation improve, becoming more standardized and more pertinent to quality evaluation. Such improvement will help minimize the problem of missing data, which is currently common in patient records. Until proven otherwise, "missing" is interpreted as "not done" for both medical liability and quality review.

Training and Education

Education of the individuals and groups involved in developing and applying guideline-derived evaluation tools is vital to the successful use of these tools. Education and training issues should receive considerable attention during the entire process of selecting the multidisciplinary team and developing and using the evaluation tools.

Three types of individuals are specifically targeted for training: the developers of the evaluation tools, the data abstractors, and the users of the performance data. In some situations, representatives from all three of these groups may be included on the multidisciplinary team.

The *developers of the evaluation tools* should be familiar and experienced with the clinical aspects of the patient care they plan to evaluate. Before embarking on the guideline translation process, the group should be educated in the following areas:

- Attributes of good clinical practice guidelines.
- Attributes of good medical review criteria.
- Differing purposes of clinical practice guidelines and medical review criteria.
- Methodology for deriving medical review criteria, performance measures, and standards of quality from clinical practice guidelines.
- Differences between the use of guideline-derived evaluation tools and the implicit review process that has been customarily used (see Chapter 4, pp. 4–5).

The developers also should be familiar with the resources that are available for their quality evaluation project and understand the data required, potential data sources, data validity and quality issues, and the information systems available to retrieve and analyze the data.

Frequently, organizations employ nonclinicians to function as *data abstractors*. The educational requirements for these individuals are significantly affected by the clinical complexity and sophistication of the data source. Data abstractors should receive training in the following:

- Language and format of the medical review criteria set.
- Data element definitions.
- Data collection codes and instrument.
- Special circumstances of data collection.
- Medical terminology of the health care condition under evaluation.
- Clinical practice guideline and its translation into medical review criteria.
- Purpose of the collected data elements.
- Confidential nature of health care data and performance measurement, and the use of appropriate safeguards for confidentiality of both patients and providers.

The training of data abstractors should include pilot testing of data collection forms and procedures. At the completion of the pretest, data abstractors should be given feedback about the correctness of their data-gathering efforts. In addition, it is essential to perform ongoing data quality checks for accuracy of data abstraction. Providing feedback of this information during the review process is essential (see Chapter 5, pp. 57–60).

Members of the multidisciplinary team that develops and implements an evaluation tool may be the *primary users of the performance data*. If the

primary users of performance data are to interpret the data accurately, they too need education and training. The amount and type of training required will depend on the extent of their professional and clinical backgrounds. Effective analysis of performance rates requires that the users be knowledgeable about the processes of health care delivery and that they have received some training in statistical techniques, including data display and interpretation.

For example, performance data users will need to know about several aspects of the review criteria development and implementation process, including the following:

- The identified purpose of the evaluation.
- The clinical practice guideline used as the source.
- The data sources and data collection methodology.
- Any limitations to data collection or analysis, such as missing data (unobtainable record documentation).
- Application of performance rates in order to improve health care.
- The confidentiality of the information.

Variation in Process

There has been great variation in how evaluation tools have been developed and used (Kellie and Kelly, 1991). One reason is that past developers of evaluation tools have not necessarily disclosed the guidelines or other source documents that form the basis of their performance measures, the process by which the measures were developed, or even the measures themselves. In addition, no generic methodology has been developed to guide the process of developing review criteria. In fact, a fundamental assumption of the workgroup was that a generic methodology should be developed to guide the translating of clinical practice guidelines into evaluation tools. While the methods presented in this document are useful for deriving evaluation tools from the AHCPR-supported guidelines, the workgroup believes that these same methods are adaptable for other guidelines, to the extent that the other guidelines have been developed with similar rigor.

Issues Related to the Translation Methodology

The workgroup members and those who provided comments repeatedly raised similar issues and concerns related to the process of developing evaluation tools, including the need to be selective, being sure to evaluate the guideline itself, reviewing periodically, planning for implementation, potential barriers, and integration with ongoing quality management activities.

Need To Be Selective

Individual users must identify a topic for review and a corresponding clinical practice guideline from which to derive evaluation tools. The choice of topic depends on a number of factors, including interest and support from the health care professional staff and the prevalence of the clinical condition within the organization. One of the most important factors to consider in selecting a topic is the perception that there is a problem with the quality of care received by patients with the selected condition. In the spirit of continuous quality improvement, there is always room for improvement of processes. However, it seems appropriate to start where there is the most opportunity for improvement—that is, where there is a perceived problem with the quality of care, and where change is feasible and can be realistically made.

Once a clinical condition has been selected for evaluation, an appropriate practice guideline must be identified. The workgroup recognizes that several guidelines may be available for a particular clinical condition. However, it is not the purpose of this document to give detailed guidance on choosing the most pertinent guideline. The eight attributes of good practice guidelines, as defined and described by IOM (Table 3.2), are an important

Table 3.2. Attributes of good practice guidelines

Attribute	Discussion
Validity	Practice guidelines are valid if, when followed, they lead to the health and cost outcomes projected for them, other things being equal. A prospective assessment of validity will consider the projected health outcomes and costs of alternative courses of action, the relationship between the evidence and recommendations, the substance and quality of the scientific and clinical evidence cited, and the means used to evaluate the evidence.
Reliability/reproducibility	Practice guidelines are reliable and reproducible (1) if—given the same evidence and methods for guidelines development—another set of experts would produce essentially the same statements, and (2) if—given the same clinical circumstances—the guidelines are interpreted and applied consistently by practitioners or other appropriate parties. A prospective assessment of reliability may consider the results of independent external reviews and pretests of the guidelines.
Clinical applicability	Practice guidelines should be as inclusive of appropriately defined patient populations as scientific and clinical evidence and expert judgment permit, and they should explicitly state the populations to which statements apply.
Clinical flexibility	Practice guidelines should identify the specifically known or generally expected exceptions to their recommendations.
Clarity	Practice guidelines should use unambiguous language, define terms precisely, and use logical, easy-to-follow modes of presentation.
Multidisciplinary process	Practice guidelines should be developed by a process that includes participation by representatives of key affected groups. Participation may include serving on panels that develop guidelines, providing evidence and viewpoints to the panels, and reviewing draft guidelines.
Scheduled review	Practice guidelines should include statements about when they should be reviewed to determine whether revisions are warranted, given new clinical evidence or changing professional consensus.
Documentation	The procedures followed in developing guidelines, the participants involved, the evidence used, the assumptions and rationales accepted, and the analytic methods employed should be meticulously documented and described.

SOURCE: Institute of Medicine, Committee on Clinical Practice Guidelines. Clinical practice guidelines: directions for a new program. Field MJ, Lohr KN, editors. Washington, DC: National Academy Press, 1990.

starting point in assessing practice guidelines (IOM, 1990). Another helpful reference is the Provisional Instrument for Assessing Clinical Practice Guidelines, developed by IOM (1992).

When two or more clinical practice guidelines address the same topic, some recommendations may be identical. However, it is also likely that some recommendations will differ among the guidelines—perhaps because of differences in publication date (and therefore differences in the scientific evidence available for review) or differences in guideline development methodology (e.g., evidence-based process versus consensus process). In evaluating specific recommendations among relevant practice guidelines, the developers of review criteria should determine the basis for the conflicting recommendations and assess the guideline against the IOM attributes for good guidelines. Developers of medical review criteria should seek guidelines based on an explicit, evidence-based methodology, so that they may be confident that the underlying scientific evidence of each guideline provides valid linkages between interventions and health outcomes.

Clinical practice guidelines, whether developed by AHCPR or by other groups, vary in the amount of underlying scientific evidence. Furthermore, specific guideline recommendations may vary in strength of scientific evidence. The recommendations for clinical care in the AHCPR-supported

guidelines represent the synthesis of scientific evidence or, in the absence of scientific evidence, the expert opinion of the multidisciplinary panel. These guideline recommendations generally indicate the strength of scientific evidence used to derive the recommendation. This evidence exists on a continuum from strong scientific support (e.g., the recommendation is based on several randomized controlled trials) to no scientific evidence, (e.g., the recommendation is based on expert opinion). For example, AHCPR-supported guideline panels may classify the strength of evidence into categories such as the following:

- A—there is good research-based evidence to support the recommendation.
- B—there is fair research-based evidence to support the recommendation.
- C—the recommendation is based on expert opinion and panel consensus.

When beginning to develop evaluation tools, the developer should be aware of the type of evidence supporting the relevant guideline recommendations and the strength of evidence assigned to these recommendations.

Evaluating the Guideline Itself

There has been considerable variation in the process of developing evaluation tools. A study of the Medicare PROs has documented such variation, reporting that while some evaluation tools adhered to the practice guidelines developed by national physician organizations, others were substantially modified. In some cases, the medical review criteria bore no resemblance to the recommendations in the guidelines (Kellie and Kelly, 1991).

Yet the purpose of developing review criteria is to evaluate to what extent a practice guideline has improved quality of care. To achieve this purpose, it is important to maintain the integrity and intent of the clinical practice guideline by developing medical review criteria that mirror the recommendations of the guideline selected. The more deviation between the review criteria and the recommendations in the guideline, the less useful or appropriate the criteria in evaluating the actual impact of the guideline.

The following two integrity checkpoints are met by medical review criteria that accurately represent the clinical practice guideline from which they were derived:

1. The criteria express the intent of the guideline, and the accuracy of that relationship has been validated by clinicians using the criteria.
2. The criteria encompass all aspects of the guideline relevant to the topic under review; for example, if the guideline includes acceptable alternatives to preferred treatment decisions, the review criteria also reflect these acceptable alternatives.

Reviewing Periodically

Like other aspects of clinical care, the development and use of evaluation tools is a dynamic process. The performance measures must be reviewed and revised over time. Some practice guidelines and evaluation tools may remain essentially the same for years, while others may need to be revised and updated fairly often because of new scientific evidence. In addition, applying evaluation tools to measure compliance with the recommendations of a guideline may bring to light deficiencies in the practice guideline itself and promote revisions of it. As Figure 2.3 shows, the use of practice guidelines and the application of guideline-derived evaluation tools are interrelated. The use of outcome measures to revise and update clinical practice guidelines and guideline-derived evaluation tools is discussed in Appendix A.

Planning for Implementation

The successful use of evaluation tools requires the development of specific implementation plans and systems (Audet, Greenfield, and Field, 1990; Berwick, Godfrey, and Roessner, 1990; Goldfield and Nash, 1989; Gottleib, Margolis, and Schoenbaum, 1990; Kane and Lurie, 1992; Schoenbaum,

1990; Schoenbaum and Murrey, 1991; Winslow, 1990). The intended purposes of and uses for the evaluation tools must be delineated before they are developed. Similarly, the timeframe for the care being evaluated must be carefully determined and specified. This timeframe usually will be retrospective, but evaluation tools also may be used in prospective and concurrent review programs.

The successful application of evaluation tools requires that clinicians whose care is being evaluated "buy into" the process from the beginning. Clinician buy-in requires attention to a number of factors:

- Ensuring participation of all stakeholders in the development and implementation process.
- Achieving effective and timely communication.
- Identifying an appropriate person to "champion" performance measurement within the organization.
- Making sure that evaluation tools address topics that are meaningful to those whose performance is being evaluated.
- Ensuring that evaluation tools are based on valid, explicit clinical practice guidelines.

Potential Barriers

It is helpful to identify and plan for barriers that may be encountered in implementing evaluation tools. The following are examples of such barriers:

- Lack of buy-in to the development and the implementation of evaluation tools (Berwick, Godfrey, and Roessner, 1990; Berwick and Wald, 1990; Goldfield and Nash, 1989; Kane and Lurie, 1992; Nash, 1990), including fear of what the use of specific evaluation tools will mean, especially to clinicians.
- Resistance by health care practitioners to intrusion on their clinical autonomy (Nash, 1990).
- Economic self-interest, including the impact of incentives or disincentives that might evolve from the use of review criteria and performance measures (Hillman, Pauly, and Kerstein, 1989; Lomas, Anderson, and Dominick-Pierre, 1989).
- Lack of understanding about the methods used to develop guidelines and evaluation tools.
- Administrative costs generated by the implementation of evaluation tools (Winslow, 1990).
- Inadequate education about the purpose of the evaluation tools (Ackerman and Nash, 1991).

Integration With Ongoing Quality Management Activities

The application of guideline-derived evaluation tools will be most effective when the tools are integrated with ongoing quality management efforts. The use of evaluation tools is one way to assess quality of care (i.e., the degree of conformance to specific guideline recommendations). Most individual practitioners and provider organizations will need quality improvement measures in addition to guideline-derived evaluation tools. Therefore, the use of these evaluation tools represents only a portion of an organization's total quality improvement efforts and should be consistent with the overall quality goals of the organization.

When medical review criteria and performance measures are applied, the resulting data document a distribution of performance (Schoenbaum and Murrey, 1991). Individuals who are interested in changing clinical performance can use this information to prompt change or shift that distribution. For example, efforts may be made to (1) eliminate unacceptable performance (truncating the distribution to include acceptable performers only); (2) shift to better practice (shifting distribution of performance toward a better mean performance); or (3) reduce unnecessary variation (narrowing distribution of performance around the mean, an effect that always occurs as the mean approaches 0 or 100 percent).

The workgroup recognizes that current quality theory emphasizes the importance of continuous improvement, of embedding the pursuit of quality into day-to-day management, of *building* quality in rather than *inspecting* it in (Berwick, Godfrey, and Roessner, 1990; Kane and Lurie, 1992; Leatherman, Peterson, Heinen et al., 1991). In general, guideline-derived evaluation tools should be considered a way to support activities that seek to identify opportunities for improvement in the delivery of health care. This approach involves data feedback to all the users of quality-related data (Berwick, Godfrey, and Roessner, 1990; Goldfield and Nash, 1989; Kane and Lurie, 1992; Leatherman, Peterson, Heinen et al., 1991; Schoenbaum and Murrey, 1991).

Possible Impact on Health Care Practitioners

A Tool, Not a Rule

Some practitioners fear that evaluation tools may be used to compile evidence of nonconformance to "standards of care" and will become part of a punitive process. Most quality assurance efforts in the past have focused on identifying "bad apples" and in some way punishing them. Research has documented that this approach is inadequate and that poor outcomes are far more often due to failures in the processes of care, not to individual error (Bader, 1992). Guideline-derived evaluation tools are not intended to provide data that will promote slavish adherence to a specific and rigid way of providing care, but to serve as a mechanism to evaluate overall quality of care.

Guideline-derived performance measures are tools for providing data related to quality of care. How the data obtained from applying these tools are used should depend on what is known about the clinical practice guideline, the accuracy and reliability of the data, and the confidence intervals for the data. In some cases, it may be possible to reach 100-percent compliance with guideline recommendations; in other cases it may not. Guidelines and the evaluation tools derived from them allow for clinical judgment and for patients to select among alternative treatments according to personal preference.

Standards of Quality

One major issue related to standards of quality is that a common understanding of this term be reached. As defined by the IOM and the workgroup, standards of quality are "authoritative statements of (1) minimum levels of acceptable performance or results; (2) excellent levels of performance or results; or (3) the range of acceptable performance or results" (see Table 2.1).

The development and application of standards of quality raises several issues for those involved in providing health care services. The principal issues are concerns that standards not be arbitrary and that any actions, especially punitive actions, prompted by care falling outside these standards be clearly warranted.

An additional paramount issue is the matter of who should determine standards of quality for guideline-derived evaluation tools. The workgroup believes that societies of health care professionals should provide guidance in setting these authoritative statements.

Process Versus Outcome Measurement

With increasing emphasis on the outcomes of care, clinicians want assurance that the standards of quality used for patient outcomes employ a methodological approach that appropriately adjusts the results for such patient characteristics as age, sex, case mix, and severity of illness. The level of analytical sophistication necessary to develop appropriate outcome performance measures and standards may be significantly greater than the sophistication needed to determine whether care processes conform to clinical practice guidelines. For more discussion of outcome measurement, see Appendix A.

Public Disclosure

The implementation of an institutionwide quality evaluation may bring about unexpected results. Projects designed to monitor individual clinicians' patterns of care may suddenly gain the attention of the payer and media communities. Also, individual patients and consumer groups may request access to data or action plans. Whether such quality review information is protected from public access may vary from State to State. It is important to determine what data are protected before implementing the evaluation tools in a quality improvement program. Health care organizations are being asked to provide increasing amounts of data to document the quality and value of care they provide, including evidence of compliance with clinical practice guidelines. Therefore, it is recommended that a spokesperson be identified to represent the organization and respond to requests for quality-related information. A policy and procedure statement will also help guide internal access to and use of the data collected for quality evaluation purposes.

Organizations opting to participate in a consortium face additional issues related to data pooling; blinding of data; and the comparative analysis, general access, and use of data, as well as publication rights. It is critical to consider these issues before implementing a quality evaluation process and to develop a consensus that is reflected in written policy and procedure statements.

Legal Issues

Several legal liability issues are raised by using clinical practice guidelines as the basis for developing instruments to evaluate and improve the quality of care. Perhaps the most prevalent legal issue currently receiving attention is the potential use of such instruments in resolving medical malpractice claims. It seems logical to speculate that rigorous methodological documentation of the most scientifically efficacious approaches to treatment will ultimately improve the overall quality of care, thereby decreasing the incidence of malpractice and protecting doctors and hospitals against malpractice suits. However, there is a paucity of evidence to confirm this supposition. Many legal experts believe it is equally possible that guidelines could make it easier to either prove or defend a case—that is, guidelines could be both a sword and a shield.

Guideline-derived evaluation tools raise other questions of legal liability as well. Developers of guidelines and evaluation tools are concerned that they could be sued by injured patients claiming the resulting quality criteria lowered the standard of care. In addition, there are fears of suit by medical suppliers whose products are not recommended by the guideline or the guideline-derived review criteria. Further, insurers worry that if they deny reimbursement for services that guidelines indicate are ineffective, inappropriate, or unnecessary, they could be subject to "bad faith" claims.

At present, it is speculative at best to predict the legal effects of evaluation tools based on guidelines. Courts may be suspicious of guideline-derived standards of quality, performance measures, or medical review criteria believed to have been promulgated by specific groups for self-protection interests or motivated primarily by cost containment efforts. On the other hand, if courts view guidelines and the resulting evaluation tools as legitimately encouraging high-quality, effective, and appropriate health care delivery based on the most current scientific knowledge, they may give considerable deference to guidelines and evaluation tools. Similarly, courts may consider these same factors when faced with questions regarding the liability of guideline developers and those who translate guidelines into evaluation tools.

Because the usefulness of guideline-derived evaluation tools depends largely on the scientific and clinical soundness of the guidelines, workgroup members strongly support full disclosure of review criteria. Such disclosure includes information underlying the development of clinical practice guidelines as well as the explicit rationale and methodology used in translating guidelines into evaluation tools. Many experts believe that disclosing review criteria as well as information about mortality, morbidity, and complication rates will

empower consumers to play a larger role in medical decisionmaking and make them more risk adverse. Patients who understand the risks and benefits inherent in various treatment options and who participate fully in the decisionmaking process regarding their care may be less prone to challenging quality and payment review decisions. Similarly, if guideline-derived review criteria are to succeed in changing provider behavior toward improving quality, providers must understand prospectively the criteria that will be used to evaluate the quality, effectiveness, and appropriateness of the care they deliver.

The uncertainty surrounding liability issues in clinical practice guideline development and translation of guidelines into evaluation tools is not likely to become explicit in the immediate future. The potential legal implications are as numerous as the current variety of purposes, approaches, and methods for developing guidelines. Organizations and institutions planning to engage in translating guidelines into evaluation tools may wish to obtain legal counsel knowledgeable in the field.

Summary

The issues discussed above represent a broad range of ideas and concerns that were raised in the process of developing recommendations for using clinical practice guidelines as quality measures. Many of these concerns were expressed by workgroup members as well as representatives from the public who provided oral and written testimony at the open meeting. Wise use of resources is recommended, as is careful integration of this new activity with already established systems for quality assessment or improvement. However, while consideration should be given to each of these issues, they should not dissuade individuals or organizations from proceeding to develop and use guideline-derived evaluation tools as one component of an overall quality improvement effort.

References

Ackerman FK III, Nash DB. Teaching the tenets of quality: a survey of medical schools and programs in health administration. QRB 1991;17:200–3.

Audet A-M, Greenfield S, Field M. Medical practice guidelines: current activities and future directions. Ann Intern Med 1990;113:709–14.

Bader BS. Rediscovering quality. Rockville, MD: Bader and Associates; 1992.

Berwick DM, Godfrey AB, Roessner J. Curing health care. San Francisco: Jossey-Bass Publishers; 1990.

Berwick DM, Wald DL. Hospital leaders' opinions of the HCFA mortality data. JAMA 1990;263:247–9.

Diokno A, McCormick KA, Colling J., et al. Urinary incontinence in adults. Clinical Practice Guideline. AHCPR Pub. No. 92–0038. Rockville, MD: Agency for Health Care Policy and Research, Public Health Service, U.S. Department of Health and Human Services; March 1992.

Goldfield N, Nash DB, editors. Providing quality care: the challenge to clinicians. Philadelphia: American College of Physicians; 1989.

Gottlieb LK, Margolis CZ, Schoenbaum SC. Clinical practice guidelines at an HMO: development and implementation in a quality improvement model. QRB 1990;16:80–6.

Hillman AL, Pauly MV, Kerstein JJ. How do financial incentives affect physicians' clinical decisions and the financial performance of health maintenance organizations? N Engl J Med 1989;321:86–92.

Institute of Medicine (IOM), Committee on Clinical Practice Guidelines. Clinical practice guidelines: directions for a new program. Field MJ, Lohr KN, editors. Washington, DC: National Academy Press; 1990.

Institute of Medicine (IOM), Committee on Clinical Practice Guidelines. Clinical practice guidelines: from development to use. Field MJ, Lohr KN, editors. Washington, DC: National Academy Press; 1992.

Kane RL, Lurie N. Appropriate effectiveness: a tale of carts and horses. QRB 1992;18:322–6.

Kellie SE, Kelly JT. Medicare Peer Review Organization preprocedure review criteria. An analysis of criteria for three procedures. JAMA 1991;265:1265–70.

Leatherman S, Peterson E, Heinen L, Quam L. Quality screening and management using claims data in a managed care setting. QRB 1991;17:349–59.

Lomas J, Anderson GM, Dominick-Pierre K. Do practice guidelines guide practice? The effect of a consensus statement on the practice of physicians. N Engl J Med 1989;321:1306–11.

Nash DB. Practice guidelines and outcomes. Where are we headed? Arch Pathol Lab Med 1990;114:1122–5.

Schoenbaum SC. Implementation of preventive services in an HMO practice. J Gen Intern Med 1990;5:S123–7.

Schoenbaum SC, Murrey KO. Impact of profiles on medical practice. Background paper for conference on profiling, sponsored by Physician Payment Review Commission. Washington, DC; January 15, 1991.

Winslow C. The role of guidelines in achieving rational health care management. Internist 1990;May;14–6.

Appendix A. Outcomes as Performance Measures[1]

Interest in Outcomes Measures

Patient health outcomes (e.g., morbidity and mortality rates, surgical site infection rates, cancer survival rates) have been measured for many years. Recently, the use of patient health outcomes as performance measures has steadily increased. Some have seen patient health outcomes as the gold standard for measuring performance in the provision of health care. Good patient health outcomes *are* the ultimate goal of all health care services. And, when health care outcomes are compared with the costs of health care services used to achieve them, patients, payers, and others make judgments about the *value* of the health care services.

However, the proper use of outcomes to measure performance in health care depends on understanding three points:

1. The limitations in collecting and analyzing outcomes data.

2. The factors that must be considered in interpreting and using outcomes data.

3. The fact that patient health outcomes achieved on one occasion by a provider organization or a clinician will not help predict future outcomes unless the processes that led to those outcomes are stable (e.g., conform to a clinical practice guideline).

The patient health outcomes of primary interest in health care include the following:

- Prevention of preventable illness, disability, and death.

- Restoration (or maintenance) of function affected by disease, including

  - physical function, such as locomotion, vision, or activities of daily living, including feeding and toileting;

  - psychological function, such as memory, cognition, or mood;

  - social function, such as interpersonal relations or communication; and

  - role function, such as that of a parent, a student, or a worker.

- Cure (or retardation) of disease, such as infection, cancer, or hypertension.

- Relief of discomfort, including

  - physical discomfort, such as pain or paresthesias; and

  - psychological discomfort, such as depression.

Any of these patient health outcomes can be measured for an individual patient and, by combining the outcomes of many individuals, for groups of patients or for general populations. The measurement of outcomes begins at the individual patient level, and these measurements can be used to assess the effectiveness of the treatment plan for the

[1]*Author: Paul M. Schyve, M.D.*

patient and to modify the plan on the basis of the outcomes. The outcomes from many individual patients can then be pooled to assess factors of interest across patients.

Such pooled outcomes measures play a central role in the development, assessment, use, and revision of clinical practice guidelines. Clinical practice guidelines describe selected processes of direct care (e.g., how to treat benign prostatic hyperplasia, how to relieve pain after surgery) and the support processes around them, such as the use of diagnostic imaging and laboratory tests, that are believed—on the basis of varying degrees of empirical evidence—to increase the likelihood of good patient outcomes. With respect to these practice guidelines, the following can be said of outcomes measures:

- Outcomes measures can be used to assess the effectiveness of alternative interventions (e.g., medication vs. surgery for benign prostatic hyperplasia, various medications to reduce the risk of repeat myocardial infarction) or alternative practices (e.g., the surgical technique for a cholecystectomy) in developing and revising clinical practice guidelines and guideline-derived evaluation instruments.

- Outcomes measures, when compared to measures of the degree to which the recommendations in existing clinical practice guidelines are implemented, can be used to evaluate the effectiveness of the recommendations.

- Outcomes measures can sometimes substitute for direct measures of the frequency with which recommendations in a clinical practice guideline are followed (i.e., the degree to which the guideline is implemented) when measuring the process of care is especially difficult or costly. For example, measuring patients' pain leveis after surgery—an outcome—may substitute for measuring whether each of the recommended steps in pain management guidelines is followed in practice (assuming the guidelines are effective).

A desirable performance measurement system includes measures both of patient health outcomes and of the degree to which the processes that led to these outcomes are stable and conform to scientifically based knowledge (as reflected, for example in clinical practice guidelines). While the assessment of outcomes measures can provide information about the past and current level of performance, the assessment of process measures can provide information about the sustainability of that performance and about the relationship of the processes to the outcomes.

Other examples of the pooling of outcomes measures include the following:

- The outcomes of patients with a specific condition (e.g., rectal cancer, osteoarthritis, schizophrenia) can be pooled to determine the usual course of the disease, including variations, and the impact of current interventions.

- The outcomes of patients who are at risk of poor outcomes (e.g., patients with angina) can be pooled to assess the impact of various risk factors (e.g., diet, exercise) on the outcomes.

- The outcomes of patients receiving a specific intervention (e.g., total hip replacement, thrombolytic therapy in acute myocardial infarction, coronary artery bypass graft) can be pooled to assess the usual results of the intervention, including variations in results.

- The outcomes of patients receiving treatment from a specific health care organization (or an individual clinician) can be pooled to assess the results achieved by the organization (or the clinician).

- The outcomes—especially in terms of incidence and prevalence of disease, disability, and death—for a population, such as a geographic area (e.g., city, State) or the employees of an employer, can be surveyed to assess the effect of the health care delivery system, the population's lifestyle (e.g., diet), and envi-

ronmental factors (e.g., environmental pollution) on the population's health status.

Because patient health outcomes have these numerous uses, especially when pooled across many patients, measuring them is of interest to various individuals, groups, and organizations. Patients, of course, are intensely interested in the outcomes of their own treatment and should participate in determining the full spectrum of outcomes of their care that are to be measured. For example, patients with cancer may be just as interested in outcomes related to the quality of life they will experience (i.e., physical, psychological, social, and role functioning, and freedom from pain) as they are with their 5-year survival after treatment. Both provider organizations and individual clinicians are interested in outcomes, as are payers, utilization management organizations, regulatory agencies, certifying and accrediting bodies, population health authorities (e.g., city and State public health authorities), health services and clinical researchers, and health policymakers. Possible purposes for which these various individuals, groups, and organizations use outcomes measures are described below.

Relationship of Outcomes to Performance

Although good patient health outcomes are the ultimate goal of health care services, the relationship between these outcomes and the performance of a health care organization (or of an individual clinician) is complex. The activities of provider organizations and of clinicians—what they do and how well they do it—are certainly causes of patient health outcomes. But they are only partial causes, and it is often impossible to definitively attribute a specific outcome—whether good or bad—to the performance of a provider organization or an individual clinician. Other patient-related and community-related factors, both known and currently unknown, will affect patient health outcomes. To relate performance to an outcome, one must consider the temporal relationship between the two, the ability to measure the outcome reliably, and the other factors that may affect the outcome.

The closer a measured outcome is in time to the activities of the provider organization or clinician whose performance is being evaluated, the more likely that the level of performance has at least partly contributed to the outcome; the longer the elapsed time between the treatment and the measurement of the outcome, the more likely that other patient-related factors, community-related factors, or even factors related to other provider organizations or clinicians will have affected the outcome. Consequently, it can be useful to classify outcomes into three temporal groups—immediate, intermediate, and long-term:

- *Immediate* outcomes emerge during or immediately after a step or specific procedure in the care process (e.g., bleeding from a complication during a surgical procedure; an adverse drug-drug interaction after administration of a medication). Immediate outcomes can frequently be attributed to the specific activity during which they occurred or that immediately preceded them.

- *Intermediate* outcomes occur at the time an episode of care is completed and can usually be associated with the procedures and other activities during that episode (e.g., the severity of cardiac and respiratory distress at discharge from a hospitalization for congestive heart failure).

- *Long-term* outcomes are the ultimate outcomes of the treatment for a particular disease or condition (e.g., the level of pain-free physical mobility 2 years after a total hip replacement, the 5-year survival rate for cancer). Often, long-term outcomes are of special interest to health services and clinical researchers, health policymakers, and population health authorities as well as to patients.

Agreement on the desirability of measuring certain outcomes may come quickly, but measuring those outcomes accurately may prove difficult. For example, anatomic and physiological variables (e.g., blood pressure, range of motion of a joint) are generally easier to measure than functional levels

(e.g., ability to perform physical work, social functioning). And intermediate outcomes (e.g., degree of freedom from pain at discharge after a total hip replacement) generally can be more reliably measured than long-term outcomes (e.g., functional mobility 2 years after hip replacement) because of the difficulties in following ambulatory patients over time.

Factors related to the patient, the clinician, the provider organization, and the community affect outcomes. When outcome measures are used to assess the performance of a health care organization or a clinician, the factors not related to the organization's or clinician's performance must be considered. For example, in evaluating the relationship between a patient's outcomes and the performance of a surgeon in a total hip replacement procedure, factors related to the patient, to other clinicians who were involved in the patient's treatment (e.g., anesthesiologist, nurses, physical therapists), to the organization (e.g., the equipment in the operating suite and nursing units), and to the community (e.g., the availability of physical therapy services) must be considered.

Factors That Affect Outcomes

Patient-related factors that affect outcomes include the following:

- Normal biological variations among patients (e.g., those associated with sex and age).
- Severity of the patient's illness.
- The patient's other illnesses (comorbidities).
- The patient's preference (and consent) for a specific treatment intervention or lack of intervention.
- The patient's involvement in the treatment.
- Certain other patient-related factors (e.g., socioeconomic level, education, occupation, environment, home, family, payment source).

The clinician-related factors that affect outcomes include the following:

- Clinician's level of knowledge (based on education and experience).
- Clinician's current technical and interpersonal skills, including empathy (based on personality, training, and experience).
- Clinician's judgment (based on education, training, and experience).

The provider organization-related factors that affect outcomes include the following:

- The availability of the following, and how they are managed: human resources (e.g., numbers and qualifications of staff, continuing education); capital resources (e.g., space, equipment); information resources (e.g., patient health records, library, ethics consultation); and financial resources.
- The organization's systems and processes for governance, management, clinical care, and support services.

The community-related factors (i.e., factors in the provider organization's external environment) that affect outcomes include the following:

- The structure of the community's health care system (e.g., whether it is integrated to provide continuity of care across provider organizations and clinicians, across organizations, and through time).
- The amount and allocation of health care resources in the community (e.g., financial resources available for health care, allocation of resources between preventive and treatment services, shortages of nursing staff).
- Other factors that affect the population's health (e.g., environmental pollution, workplace safety, socioeconomic and educational levels).

In establishing a strong possible causal link between the performance of an organization or of a clinician and a measured outcome, the contribution of these other factors to the outcome must be considered. Sometimes it is useful to evaluate the raw outcomes data and then retrospectively consider these other factors in explaining any possible relationship between the performance and the outcomes. Often it is more useful to prospectively control for the contributions of these other factors; that is, their contributions must be eliminated or held constant, either in reality or statistically. Since these other factors cannot usually be eliminated, an attempt can be made to hold their contribution constant through one of four methods (or a combination thereof). The four methods are as follows:

1. Select the patients whose outcomes are to be measured in such a way that the other factors that affect the outcomes are held constant (e.g., measure the outcomes only of hypertensive patients who have the same severity of illness). The large number of factors that would have to be held constant by this method may make controlling for all of them impractical.

2. Stratify the selection of patients whose outcomes are measured so that they are studied in groups in which the other factors that affect the outcomes are held constant (e.g., measuring outcomes separately for acute myocardial infarction patients with and without a specific comorbidity, such as diabetes). Again, because the number of factors that need to be controlled is so large, the patients may need to be stratified into many groups, such that each group may be too small to study statistically.

3. Measure changes in health outcomes (i.e., measure the difference between the patient's baseline health status and the health status after intervention) to adjust for varying levels of severity of illness among patients; other factors are not adjusted for by this method.

4. Statistically adjust the rates of outcomes for groups of patients in which the other factors that affect outcomes are not constant, so that in any analyses they are made to act as if they are constant. For example, if the preexisting presence of hypertension is known to reduce survival rates from treated acute myocardial infarction (AMI) by a certain percentage, adjust upward the observed survival rates of provider organizations or clinicians whose AMI patients have uncommonly high rates of preexisting hypertension. In most cases, statistical adjustment is combined with some stratification in one or more of the following: data collection, subject selection, or measurement of changes from baseline.

Pooled Outcomes Data

It is often difficult to attribute the outcomes of a single patient to the performance of a provider organization or of a clinician, because it is usually impossible to identify each of the other factors listed above and to assess their contributions to the patient's outcomes. The methods identified in the previous section for controlling for these other factors rely on pooling outcomes measures—and on pooling measures of these other factors—for multiple patients. Adjusted aggregate outcome data are the outcomes of many patients, adjusted statistically to control for the differences caused by the other known factors that affect the outcomes (e.g., patient severity of illness and comorbidities). When statistically analyzed, adjusted aggregate data can be more useful than single-patient data in attributing patterns of outcomes to the performance of an organization or a clinician. Using adjusted aggregate outcome data increases the likelihood that the effect of unknown factors (or known factors that are incompletely adjusted for) on outcomes may be "washed out" by the normal variation in each factor in the larger group of patients. Further, using aggregate data permits statistical analysis of the confidence of conclusions about the association of the performance and the outcomes (e.g., through the calculation of means, standard deviations, correlations, and confidence intervals).

However, the statistical analysis of adjusted aggregate data requires an adequate statistical sample, which is not always available from a single provider organization or an individual clinician, even if data are collected over an extended period. To generate an adequate statistical sample, outcome data from the patients of many provider organizations or clinicians may be pooled, or the data may be collected over an extended period, or both. In addition, collecting multiple outcome measures over time for the same patients can increase the number of measures for analysis.

Pooling of data requires that all the data be collected using both (1) identical measures of outcomes (i.e., agreed-on specific definitions for each piece of data, or "data element," to be collected); and (2) similar methodologies for data collection (e.g., for sampling).

If the data elements and the collection methodologies are not the same, the pooled data will contain measures that are thought to be measuring the same thing but in reality are not. Pooling of data collected over time (including data from multiple outcomes measures for the same patient) requires adjusting for the changes that occur over time in the factors that affect the outcomes (e.g., changes in clinicians' practice, changes in the ages of the patient group being studied).

Even when data element definitions have been established, the reliability of data submitted to the pool may vary considerably over time or among individual data collectors, provider organizations, or clinicians. Factors affecting the reliability of data, such as the accuracy and completeness with which the data are collected and the consistency of data transmission, include (1) ambiguity in or misinterpretation of data element definitions; (2) differences or errors in coding (e.g., ICD–9–CM) practices; (3) incomplete or inaccurate documentation (e.g., in the patient's health record); (4) inadequate systems for updating data; (5) variations in or absence of policies that affect data collection (e.g., who is considered an "inpatient"); and (6) the training of data collectors and coders. If the pooled data contain unreliable data, inferences derived from the data are weakened and may be misleading.

Reference Data Bases and Comparative Data

A data base of pooled reliable outcomes data, and the data needed to adjust for the other factors affecting outcomes, can be used as a comparative, or "reference," data base. The data in a reference data base can be used, with adjustment for other factors that affect outcomes, to compare the outcomes from one provider organization or clinician with outcomes from all the other provider organizations or clinicians, respectively, or to a subset thereof (a "benchmark"). For example, a reference data base can be used to compare the outcomes of one transplant center in regard to liver transplants with those of all other centers in the data base. Such comparisons can be made among procedures, among communities, and among types of patients as well as among provider organizations and among clinicians.

The desire to use outcomes comparisons is earnest and often urgent among the following:

- Researchers and clinicians wanting to choose efficacious procedures among alternative procedures for the same patient condition.
- Patients and clinicians wanting to choose the appropriate procedure among procedures that have different outcomes (e.g., cure versus pain relief) and risks. Appropriately explained outcomes data provide information patients need in making decisions about their care.
- Provider organizations and clinicians wanting to assess and improve their performance through the feedback loop of quality assurance and improvement activities.
- Developers and revisers of clinical practice guidelines wanting to identify the most efficacious procedures among alternatives and

evaluate the impact of differences in practice on the effectiveness of the guidelines.

- Patients wanting to choose the provider organizations and clinicians they will go to for care.
- Payers wanting to select the procedures they will and will not pay for.
- Payers wanting to determine which provider organizations and clinicians will be eligible for payment and which will not.
- Government regulators, peer review organizations, certifying bodies, and accrediting bodies wanting to identify provider organizations and clinicians who require focused review as part of the licensing, certification, or accreditation process.

Limitations in the Use of Outcomes Measures for Comparative Purposes

The desire to use outcomes measures from reference data bases for these comparative purposes is understandable. Many of the decisions for which the data would be helpful are currently being made on the basis of few data or poor data. There is a danger that poor data may be used or good data may be misused, and that unwarranted and incorrect conclusions may therefore be drawn about the relative performance of provider organizations or clinicians. To understand the uses and the limitations of comparative data from any reference data base, the user must understand and consider the following:

- The data's reliability.
- The various factors that may affect the outcomes.
- The adjustment techniques used to control for the effect of those other factors.
- Which of those other factors have been included and have not been included in the adjustment.
- The adequacy of the statistical sample of outcomes.
- The techniques used for statistical analysis of the aggregate outcome data.
- The statistical confidence of the results of the analysis.
- The validity of the outcome data for their intended use (i.e., the degree to which the outcome can be attributed to the procedure, provider organization, or clinician being assessed).

Another limitation is that the comparison of trends in the outcomes data over time (i.e., whether the outcomes are improving, deteriorating, or remaining the same) between provider organizations or between clinicians often reveals more about their relative performance than point-in-time comparisons would reveal.

Conclusion

Reliable, adjusted, statistically analyzed, valid patient health outcome data can be used to compare the effectiveness of alternative procedures or the performance of provider organizations or clinicians. But when patient health outcomes are used as performance measures for comparative purposes, several factors should be kept in mind:

- Interpretation of the outcomes should be based on (1) the degree of reliability of the data, (2) the adjustments that were made to the data to render them comparable, (3) the statistical analysis of the data, and (4) the degree of validity of the data for their intended use.
- For individual clinicians providing care within complex provider organizations

(e.g., hospitals), using patient health outcomes to compare the performance of clinicians who are in different organizations may have limited validity because of the many organization-related factors that affect patient outcomes but may not be included in adjustments of the data.

- Since adjustments of the data usually have not included *all* the other factors that affected the outcomes for patients, it is almost always more appropriate to use comparative outcomes data to trigger further evaluation than to reach a final conclusion about the level of an organization's or clinician's performance.

- To better predict future patient health outcomes, outcomes measures should be accompanied by measures of antecedent processes; this facilitates understanding the relationship between processes and outcomes, and the stability of the processes.

While the quantity of reliable, fully adjusted, adequately analyzed, valid outcome data currently available is limited, the potential use for such data, as they become available, is great.

For Further Information

Blumberg MS. Comments on HCFA hospital death rate statistical outliers. Health Serv Res 1987;21:715–39.

Elwood PM. Outcomes management: a technology of patient experience. N Engl J Med 1988;318:1549–56.

Epstein AM. The outcomes movement—will it get us where we want to go? N Engl J Med 1990;323:266–70.

Greenfield S. The state of outcome research: are we on target? N Engl J Med 1989;320:1142–3.

Institute of Medicine (IOM). Effectiveness and outcomes in health care: proceedings of an invitational conference by the Institute of Medicine. Heitkoff KA, Lohr KN, editors. Washington, DC: National Academy Press; 1992.

International classification of diseases, 9th revision, clinical modification. 4th ed. (DHHS Pub. No. (PHS) 91–1260.) Washington: Public Health Service; 1980.

Joint Commission on Accreditation of Healthcare Organizations (JCAHO). Primer on indicator development and application: measuring quality in health care. Oakbrook Terrace, IL: Joint Commission on Accreditation of Healthcare Organizations; 1990.

Joint Commission on Accreditation of Healthcare Organizations (JCAHO). The measurement mandate: on the road to performance improvement in health care. Oakbrook Terrace, IL: Joint Commission on Accreditation of Healthcare Organizations; 1993.

Lohr KN. Outcome measurement: concepts and questions. Inquiry 1988;25:37–50.

Schroeder SA. Outcome assessment 70 years later: are we ready? N Engl J Med 1987;316:160–2.

Tarlov AR, Ware JE, Greenfield S, et al. The medical outcomes study: an application of methods for monitoring the results of medical care. JAMA 1989;262:925–30.

Glossary

Acceptability. The degree to which health care satisfies patients.

Acceptable alternative. A common and legitimate reason for not conforming to practice guideline recommendations; for example, the clinician recommended a treatment according to the guideline, but the patient refused. Acceptable alternatives are specified explicitly when writing review criteria, whether the alternatives have been stated explicitly in the practice guideline or merely implied. "Acceptable alternative" is also the name of the status assigned to a criterion during a review if documentation is found for a defined acceptable alternative to the criterion.

Accessibility. The ease with which health care can be obtained.

Adverse event. An undesired event implying harm to a patient.

AHCPR. Agency for Health Care Policy and Research, Public Health Service (PHS), Department of Health and Human Services (DHHS).

Algorithm. A rule of procedure, or set of instructions, containing conditional logic for solving a problem or accomplishing a task. Guideline algorithms relate to recommendations for patient care (Gottlieb, Margolis, and Schoenbaum, 1990; Hadorn, McCormick, and Diokno, 1992; Margolis, 1983). Criteria algorithms concern rules for evaluating criteria conformance. Algorithms may be expressed in works only or in diagrammatic form (Margolis, 1992).

AMRRC. American Medical Review Research Center, a private, public interest, and education corporation.

Appropriateness. The probability of benefit to the patients exceeds the probability of harm.

Benchmark. A level of care set as a goal to be attained. Internal benchmarks are derived from similar processes or services within an organization; competitive benchmarks are comparisons with the best external competitors in the field; and generic benchmarks are drawn from the best performance of similar processes in other industries.

Case-based. Refers to a single patient or case.

Case mix. Distribution of a group of patients into categories reflecting differences in patients' diagnoses/conditions.

Case sampling period. The time period during which a case is considered eligible for inclusion in the denominator of a performance measure.

Case severity. A measure of intensity or gravity of a given condition or diagnosis for a patient.

Clinical practice guidelines. Systematically developed statements to assist practitioners' and patients' decisions about health care to be provided for specific clinical circumstances.

Comparative standard. A standard derived from a comparison with other prerformance rates constructed by using exactly the same performance measure, such as the prior performance of a clinician or provider, the observation of the performance of others, or the statistical analysis of group rates.

Confidence interval. An interval or range based on a random sample, for which there is a given probability (e.g., 95 percent) that the population mean is contained within that interval. For example, a study may show that a drug lowers the average blood pressure for patients in the study by 4.8 mm Hg, with the 95-percent confidence interval between 2.5 and 7.3 mm Hg. The confidence interval is used in performance measurement to indicate whether an individual rate from a performance review is considered statistically similar to or different from the group average rate, or from a performance rate selected to represent an acceptable level of care.

Confidence limits. The upper and lower boundaries of a confidence interval.

Criteria (singular: criterion). Standards or principles by which something is judged or evaluated.

Criteria set. A series of criterion statements linked together because they all apply to the same patient sample.

Criterion status. The category to which a case is assigned by application of a criterion. For example, the case that meets a criterion is assigned the status "met"; a case that meets an *acceptable alternative* to a criterion (see definition above) is assigned the status "acceptable alternative." If a criteria set incorporates branching logic (i.e., a certain criterion applies only to a defined subgroup of cases), cases not within that subgroup are assigned the status "not applicable" for that criterion. If data needed to determine whether a case met a criterion are not found in the selected data source, the status "not reviewable" is assigned. A case that does not fit any of the above categories is assigned by default to the status "not met" (i.e., the care given to the case did not conform to the practice guideline).

Data sample. The population or group of patients to whom a performance measure will be applied to assess rates of conformance to a clinical practice guideline. The denominator group.

Decision rule. Instructions for deciding how to evaluate clinical data. When abstracting data from medical records, decision rules determine whether and how a data item should be coded.

DEMPAQ project. The Project to Develop and Evaluate Methods for Promoting Ambulatory Care Quality. A demonstration project with the PRO program of computerized review systems using claims and records data for office-based practice; the computerized review provides comparisons of performance rates to clinician groups to assist them with identifying opportunities for quality improvements.

Denominator. For a performance measure, the sample of cases that will be observed to determine conformance to medical review criteria.

Denominator event/state, index event/state. The event or health state that defines a patient's eligibility for inclusion in the denominator group for a performance measurement.

Efficiency. Occurs when health care of the desired quality is produced at the lowest cost, or when health care produced at a fixed cost is of the highest quality.

Exclusion. Characteristics or conditions that make cases ineligible for review by a specific performance measure or by a specific criterion within a performance measure.

Explicit criteria. Objective criteria specified in advance as a basis for making judgments of performance.

External review. Review in which criteria and standards of judgment are developed or ratified with input from persons other than the clinician or clinician group that is being evaluated.

HCFA. Health Care Financing Administration, Department of Health and Human Services (DHHS).

HCQII. Health Care Quality Improvement Initiative. A program in the pilot testing phase in which peer review organizations (PROs) provide pattern analyses (i.e., comparisons of performance rates) to hospitals to assist them in identifying opportunities for quality improvement.

HSQB. Health Standards and Quality Bureau. The bureau within the HCFA responsible for the Peer Review Organization (PRO) program.

Implicit criteria. Criteria formed by a respected clinician who uses clinical judgment in evaluating performance; these implicit criteria remain concealed in the mind of the reviewer.

Implicit review. Review conducted using implicit criteria.

Index event/state. See **Denominator event/state.**

Indicator. A quantitative measure for monitoring clinical care.

Internal review. Review in which clinicians are involved in setting or adopting the criteria and standards by which they evaluate themselves.

JCAHO. Joint Commission on Accreditation of Healthcare Organizations.

Mean. Arithmetic average of the values of a sample variable.

Measure. An instrument or tool used for making measurements.

Measurement. The performance rate obtained by applying a measure.

Measurement error. Variation in measurements due to causes other than real differences in the attribute being measured.

Medical review criteria. Systematically developed statements that can be used to assess specific health care decisions, services, and outcomes. Each criterion derived from a guideline recommendation is used to determine whether the case being reviewed conforms to a particular recommendation in the guideline. A status is assigned to each criterion to reflect the care given.

Numerator. For a performance measure, the cases in the denominator group that experience events specified in a medical review criterion as evidence of guideline conformance.

Outcome data. Data describing a patient's health status.

Patterns of care. Among a group of clinicians, the distribution of the clinicians' rates of performance.

Peer review. Review conducted by a peer (a similarly qualified clinician) or peers; historically, *peer review* has been conducted by *case-based implicit review,* and so the terms are sometimes used interchangeably.

Performance indicators, indicators. Quantitative measures used to measure and improve performance and quality (JCAHO, in press). *Rate-based* indicators are similar to the *performance measures* defined in Table 2.1; they produce rates for comparing the performance of organizational providers of care. *Sentinel event* indicators identify undesired events such as death; a single instance of a sentinel event triggers a quality review (JCAHO, 1990; p.11; JCAHO, 1991, p. 43). In referring to rate-based indicators, this text uses the term *performance measures,* as does the legislation that established the AHCPR.

Performance measures. Methods or instruments to estimate or monitor the extent to which the actions of a health care practitioner or provider conform to the clinical practice guideline.

Performance rate. A measurement produced by using a performance measure, providing a quantitative evaluation of events related to patient care. A performance rate results when the numerator for a performance measure is divided by the denominator for this measure.

Population-based. The word *population* in this instance is used in the epidemiological sense—a defined group of individuals sampled for study. The term *population-based* is sometimes used to mean *rate-based.* However, a population-based measure generally is a performance rate for a group of patients in a geographic area or in a particular health plan enrollment. When used in this sense, *population-based* applies to all patients rather than to only those who use services.

Practice parameters. Strategies for patient management, developed to assist in clinical decisionmaking.

Prescriptive standard. A statement of what should be achieved rather than a statement of what has been achieved.

PRO. Peer review organization; an organization that reviews appropriateness and quality of care for beneficiaries of the Medicare program.

Process data. What is done to, for, or by patients as part of the delivery of care, such as the performance of a test or procedure.

Profiles. Sets of performance rates aggregated by clinician, clinician group, or organization to monitor some aspect of health care delivery.

Rate. A quantitative measure, usually expressed as a percentage, of the occurence of an event of interest within a specified time interval. Rates are derived by creating a fraction in which the numerator is the number of patients experiencing an event of interest and the denominator is the population of patients at risk for the occurrence of that event. A rate may also be constructed by counting events rather than patients in the numerator and denominator—that is, when the event could occur more than once for a given patient.

Reliability. The extent to which a measurement is reproducible; low levels of random error.

Sample. The subset of a population or the group of cases to whom a performance measure will be applied in order to assess rates of conformance to a clinical practice guideline.

Sensitivity. High rate of detection of "true positives."

Specificity. Low rate of detection of "false positives."

Standard for accreditation. A statement of expectation set by competent authority concerning a degree or level

of requirement, excellence, or attainment in quality or performance (JCAHO, in press).

Standard of care (legal usage). In malpractice case court proceedings there is an attempt to determine whether a patient suffered harm due to negligent violation of a standard of care. The standard of care for the case is elaborated by the questioning of expert witnesses who have studied the facts of the case before the court and have relevant knowledge of comparable behavior.

Standard of care (regulatory usage). Standards for facilities are commonly expressed in terms of a minimal level of policy, equipment, and capacity necessary to achieve licensure or certification.

Standards of quality. Authoritative statements of (1) minimum levels of acceptable performance or results, (2) excellent levels of performance or results, or (3) the range of acceptable performance or results.

Status, criterion. See **Criterion status.**

Structural data. Information about organizational facilities, equipment, policies, and procedures; for example, a hospital policy for patient-controlled analgesia.

Structured implicit review. Implicit review conducted with instructions directing the reviewer to focus on certain types of data or answer certain questions in the review process.

Technical quality. Coordination of judgment, skill, and knowledge in delivering appropriate technology to improve the health of patients.

Threshold. A preestablished level for care. If a desired attribute of care falls below this level or an undesired attribute of care rises above this level, further evaluation or action is triggered.

Timeliness. Services are completed in a timeframe that maximizes health benefit and satisfaction of the patient.

Time window. The time period following an index event during which a case is "observed" for evidence that the care did or did not conform to medical review criteria. In other words, the interval in which it must be determined whether or not a numerator event took place.

Total quality management. A management theory of quality improvement based on (1) involving the total organization, (2) using statistical quality control, (3) seeking to raise the average performance rather than eliminate outliers, and (4) continuously reevaluating performance after interventions in order to plan further interventions if needed.

Unit of analysis. The unit to which a performance measure is applied; the unit may be the patient, clinician, group of clinicians, or institution.

Validity. What is measured, how well it is measured, the effectiveness of a measure in achieving a specific purpose. *Content validity 1* is the degree to which all items in the measure relate to the performance being measured (measure is pure). *Content validity 2* is the extent to which all relevant aspects of performance are covered (measure is complete). *Face validity* is the extent to which a measure appears to measure what it is intended to measure.

Variable. A characteristic that is measured. An *independent variable* is a characteristic that explains another variable. A *dependent variable* is a characteristic that is explained by one or more other variables.

Variation. For performance rates, *variation* refers to differences between the performance rate of one clinician or group of clinicians or organizations and the performance rates of comparable others.

Peer Reviewers[1]

Arja P. Adair, Jr.
Executive Director
Colorado Foundation for Medical Care
Denver, CO

Richard F. Afable, M.D., M.P.H.
Assistant Professor of Medicine
Bowman Gray School of Medicine of Wake Forest University
Winston-Salem, NC

Mary Anne Bacas, R.R.A.
Director of Quality Assessment
Memorial Hospital
York, PA

David Ballard, M.D., Ph.D.
Director
Thomas Jefferson Health Policy Institute
Charlottesville, VA

Daniel M. Barr III, M.D.
Consultant
The MacCullouch Company
Chicago, IL

Chuck Biddle, C.R.N.A., Ph.D.
Assistant Professor of Anesthesiology
Dartmouth–Hitchcock Medical Center
Lebanon, NH

Stephen C. Biondi
Vice President, Quality Assurance and Clinical Services
Unicare Health Facilities
Milwaukee, WI

Catherine Borbas, Ph.D., M.P.H.
Executive Director
Healthcare Education & Research Foundation, Inc.
St. Paul, MN

Troyen A. Brennan, M.D., J.D., M.P.H.
Professor of Law and Public Health
Harvard School of Public Health
Boston, MA

Jerri Bryant, R.N., M.P.H.
Clinical Epidemiologist
Cleveland Clinic Foundation
Cleveland, OH

Richard E. Burney, M.D.
Professor of Surgery
University of Michigan
Ann Arbor, MI

Janet B. Chermak
Vice President of Professional Services
Hillhaven Corporation
Tacoma, WA

Kathleen Ciccone, R.N., M.B.A.
Vice President, Quality Assurance
Hospital Association of New York State
Albany, NY

Kathryn L. Coltin
Director, Clinical Management Information Center
Harvard Community Health Plan
Brookline, MA

Michel Dagher, D.O., M.B.A.
Medical Director, Emergency Department
Memorial Medical Center
Jacksonville, FL

Sidney T. Dana, M.D.
Clinical Professor
SUNY Upstate Medical Center
Syracuse, NY

[1] The individuals listed participated in peer review of a draft version of this document during the summer of 1993. Being listed in this section does not necessarily imply endorsement of this document.

Susan Dean-Baar, Ph.D., R.N.
Assistant Professor
University of Wisconsin–Milwaukee School of Nursing
Milwaukee, WI

Martha Phelps DeMers, R.R.A., M.P.A.
Quality Advisor
Keller Army Community Hospital
West Point, NY

Barbara Demster, M.S., R.R.A.
Vice President
Bottomley & Associates
Gaithersburg, MD

Louis Diamond, M.D.
Georgetown School of Medicine
Washington, DC

Betty T. Dixon, R.N., B.S.N.
Quality Improvement Coordinator
Satilla Regional Medical Center
Waycross, GA

Donalda Ellek
Manager, Office of Quality Assurance
American Dental Association
Chicago, IL

Madelon L. Finkel, Ph.D.
Clinical Professor
Cornell University Medical Center
New York, NY

Kathleen Frawley, J.D., M.S., R.R.A.
Director, Washington, DC Office
American Health Information Management Association
Washington, DC

Peter Goldschmidt
President
World Development Group
Bethesda, MD

Lawrence Gottlieb, M.D., M.P.P.
Associate Medical Director
Harvard Community Health Plan
Brookline, MA

Michael K. Greenberg, M.D.
Member, Quality Standards Subcommittee
American Academy of Neurology
Minneapolis, MN

John F. Griffin, M.D.
Madison–Irving Medical Center
Syracuse, NY

Peter A. Gross, M.D.
Director, Department of Medicine
Hackensack Medical Center
Hackensack, NJ

David Hadorn, M.D.
Consultant
RAND
Santa Monica, CA

Joette Hanna, M.P.A., R.R.A.
Data Manager, Center for Quality Resource Management
Baylor University Medical Center
Dallas, TX

J. Susan Hines
Vice President of Clinical Services and Specialty Programs
Health Care & Retirement Corporation
Toledo, OH

Susie Hutchings, R.N., B.S.N.
Nurse Consultant
Life Care Centers of America
Cleveland, TN

Barbara Irvine, P.A.
Director of Review Services
Oregon Medical Professional Review Organization
Portland, OR

Robert H. Jones, M.D.
Professor of Surgery
Duke University Medical Center
Durham, NC

Douglas G. Kelling, Jr., M.D.
Assistant Professor of Medicine
Duke University
Durham, NC

Jane N. Kimball, R.N.
Director, Nevada Operations
Nevada Peer Review
Las Vegas, NV

Kathy Kunselman, R.N.
Director, Quality Improvement
HealthAmerica of Pittsburgh
Pittsburgh, PA

Judith Lenhart, R.N.
Director, Review Operations
Colorado Foundation for Medical Care
Denver, CO

Steve Levenson, M.D.
Medical Director
Asbury Methodist Village
Gaithersburg, MD

Carole J. Magoffin, M.S.
Executive Director
American Medical Review Research Center
Washington, DC

J. Peter Maselli, M.D.
Medical Director
MassPRO
Waltham, MA

William C. Morgan, M.D.
Medical Advisor to QA/UR Department
Sarasota Memorial Hospital
Sarasota, FL

Alan R. Nelson, M.D.
Executive Vice President
American Society of Internal Medicine
Washington, DC

Karen S. O'Connor, M.A., R.N.
Director of Practice, Economics, and Policy
American Nurses Association
Washington, DC

Denis M. O'Day, M.D.
Chair, Department of Ophthalmology and Visual Sciences
Vanderbilt University School of Medicine
Nashville, TN

Delwin K. Ohrt, M.D.
Vice President/Medical Director
Blue Cross/Blue Shield of Minnesota
St. Paul, MN

Patricia A. Pickering, M.S.M., R.N., C.P.N.P.
Quality Management Specialist
Healthnet, Inc.
Indianapolis, IN

Julia A. Powell, B.S.N., M.A., C.N.A.
Vice President, Patient Services
National HealthCorp. L.P.
Murfreesboro, TN

Richard G. Roberts, M.D., J.D.
Associate Professor, Department of Family Medicine
University of Wisconsin
Madison, Wisconsin

Henry D. Royal, M.D.
Chairman, Office of Health Care Policy
Society of Nuclear Medicine
Elmwood Park, NJ

Robert Sebring, Ph.D.
Director, Division of Quality Care
American Academy of Pediatrics
Elk Grove Village, IL

Earl E. Smith III, M.D.
American College of Emergency Physicians
Dallas, Texas

L. Kent Smith, M.D., M.P.H.
Medical Director, Preventive Medicine Program
Arizona Heart Institute
Phoenix, AZ

Carl E. Speicher, M.D.
Professor and Director, Clinical Services
Ohio State University Hospitals
Columbus, OH

Marilyn P. Verhey, Ph.D., R.N.
Assistant Professor
San Francisco State University
San Francisco, CA

Jonathan Warren, M.D.
Chairman, Guidelines Committee
American College of Critical Care Medicine
Anaheim, CA

Rhonda Whitson, R.R.A.
Project Administrator, Practice Management Department
American College of Emergency Physicians
Dallas, TX

Billie Ann Young, M.B.A., R.R.A.
Chairman-Elect, Quality Assurance Section
American Health Information Management Association
Boca Raton, FL

Wanda W. Young, Sc.D.
President
The Pittsburgh Research Institute
Pittsburgh, PA

Les Zendle, M.D.
Associate Medical Director
Kaiser Permanente
Pasadena, CA

*U.S. G.P.O.:1995-380-939

http://nihlibrary.nih.gov

10 Center Drive
Bethesda, MD 20892-1150
301-496-1080

NIH LIBRARY
3 1496 00623 8532

MAY - 8 1995

U.S. Department of Health and Human Services
Public Health Service
Agency for Health Care Policy and Research
Executive Office Center, Suite 501
2101 East Jefferson Street
Rockville, MD 20852

AHCPR Publication No. 95-0045
March 1995